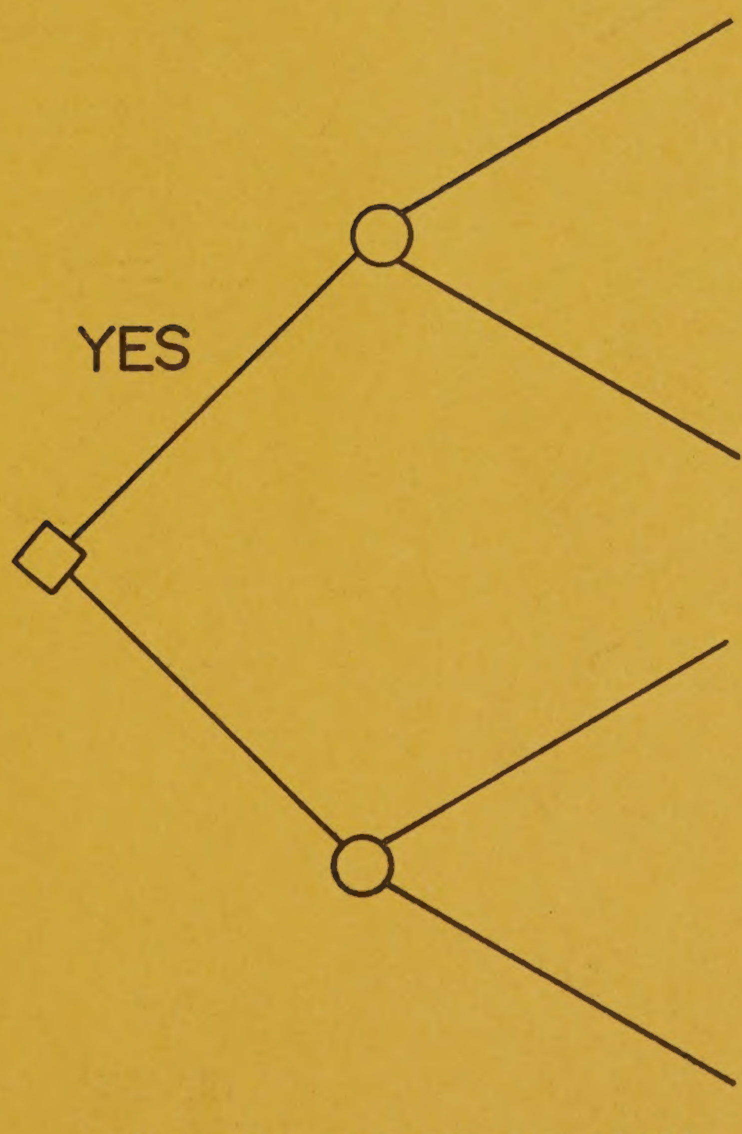
ANALYSIS OF DECISIONS UNDER UNCERTAINTY

A PRIMER

DIVISION OF ENGINEERING

U. S. FOREST SERVICE



MANUSCRIPT EDITION JUNE 1974

PREFACE

The purpose of this report is to present a technique of engineering decision making. The technique will employ the engineer's expert opinions in the form of probabilities concerning unknown conditions, and his knowledge of costs or values if these possible conditions were to exist.

Part I of this report deals with Posterior Analysis, which is selection of an action decision where all opinions and investigation results are available, and the probabilities of unknown conditions have been evaluated with this information.

Part II deals with Pre-Posterior Analysis in which the engineer feels that he does not yet have enough information to select an action, and therefore he wishes to choose a level of investigation to provide results which will better enable him to evaluate the probabilities of unknown conditions. The first example in Part II (extension of the umbrella problem) serves to introduce the concepts of Pre-Posterior Analysis. This example is followed by a sophisticated problem, the solution of which illustrates a formal and somewhat general procedure, and thus is not essential in a first reading.

"Manuscript" Statement

This edition is considered a rough draft and subject to correction before final incorporation into a Forest Service Technical Report.

This edition will be revised about January 1975 for final incorporation into the Forest Service technical report system. Interested parties should advise the Division of Engineering at the Forest Service National Headquarters of any changes they would like to see made in this document. Such changes will be considered during final revision.

We believe that techniques discussed in this edition are accurate, reliable, and the best available for the use listed. Therefore, a field unit should not hesitate to use the techniques.

We expect that changes will generally be in the area of clarification or in correcting obvious errors, rather than in changing the techniques.

ACKNOWLEDGEMENTS

This handbook was prepared from research and written material developed by the United States Forest Service Transportation Analysis Group and engineers from California State University, San Jose.

TABLE OF CONTENTS

Note: The sections indicated with * are those most essential for clear understanding of technique.

I. DECISIONS CONCERNING THE SELECTION OF AN ACTION IN THE FACE OF UNCERTAINTY (POSTERIOR ANALYSIS)

The necessity of making decisions in the face of uncertainty occurs with disturbing regularity in everyone's personal and professional life. The engineer, in particular, must make decisions which involve significant economic consequences with respect to project planning and selection of investigations related to future projects. He must employ all of his professional knowledge concerning techniques of forecasting, reconnaissance, and exploration in order to reduce the uncertainties related to future requirements, and his knowledge must also provide workable estimates of the costs and benefits related to any proposed project. He must then employ all of this knowledge to select the best type or size of project. A pertinent example involves the choice of the reliability level of flood protection for a given area: the engineer must decide whether a 90 percent reliable project with cost of one million dollars is more advantageous than a 99 percent reliable project with cost of five million dollars, when the loss due to flooding may be one million dollars.

The purpose of this report is to present a logical method of making consistent decisions. The method will employ all of the engineer's professional knowledge concerning the probability of future or unknown conditions, and information related to costs and benefits. This method is based on the concept of minimizing the expected loss, or maximizing the expected gain, and can be introduced by considering the following situation.

Let us suppose that we are invited to indulge in the following game of chance: a die is rolled and if ace occurs, we win one hundred dollars, and if not ace, the we lose one dollar. Very few, if any of us would refuse this fairly good chance of winning one hundred dollars because of the relatively small amount of possible loss. Our decision to accept this attractive gamble is based on a thinking process which is mathematically equivalent to a computation of expected gain: defined as,

$$(\text{winnings}) \times (\text{chance of success}) - (\text{loss}) \times (\text{chance of failure}).$$

For example, in our game the expected gain is

$$(\$100)(1/6) - (\$1)(5/6) = \$15.84$$

It can be shown that a decision based on the maximum expected gain, or minimum expected loss, is economically sound. In engineering, economic decisions involve gambling in the face of uncertainty quite similar to our proposed game of chance, and therefore it is necessary to develop a minimum expected loss calculation procedure, involving engineering estimates of probabilities or chances and costs, for any given decision problem. This general decision procedure or method will be developed in terms of the following, very human, example in which the losses and probabilities are readily available:

A cosmopolitan gentleman has deemed it absolutely necessary to attend a theatrical performance. He must walk home from the theatre, and therefore consults the newspaper weather report for the possibility of rain at the end of the performance. The current method of weather forecasting furnishes the gentleman with a 40 percent chance for rain at the time that he must walk home. His decision in the face of this uncertainty of rain is whether or not to carry an umbrella. If he goes without the umbrella, and rain occurs, then he must pay two dollars to press his rainsoaked suit. If he elects to carry the umbrella, then he must pay a fifty-cent fee to check in the umbrella at the theatre cloakroom.

I(a) Symbols and the Decision Tree

In order to best present all of the information in this problem, and to provide a general calculation procedure in terms of this information we will define the following symbols and introduce the concept of the decision tree.

List of Decision choices or Actions {1,2}

1 = Take umbrella

2 = Do not take umbrella

List of Future Actual Conditions or Situations {A,B}

A = rain at end of theatrical performance

B = no rain at end of theatrical performance

Probabilities of Future Actual Situations P(A), P(B)

P(A) = probability of rain = 0.4

P(B) = probability of no rain = 0.6

Losses (L) corresponding to a selected action (1 or 2) and actual situation (A or B).

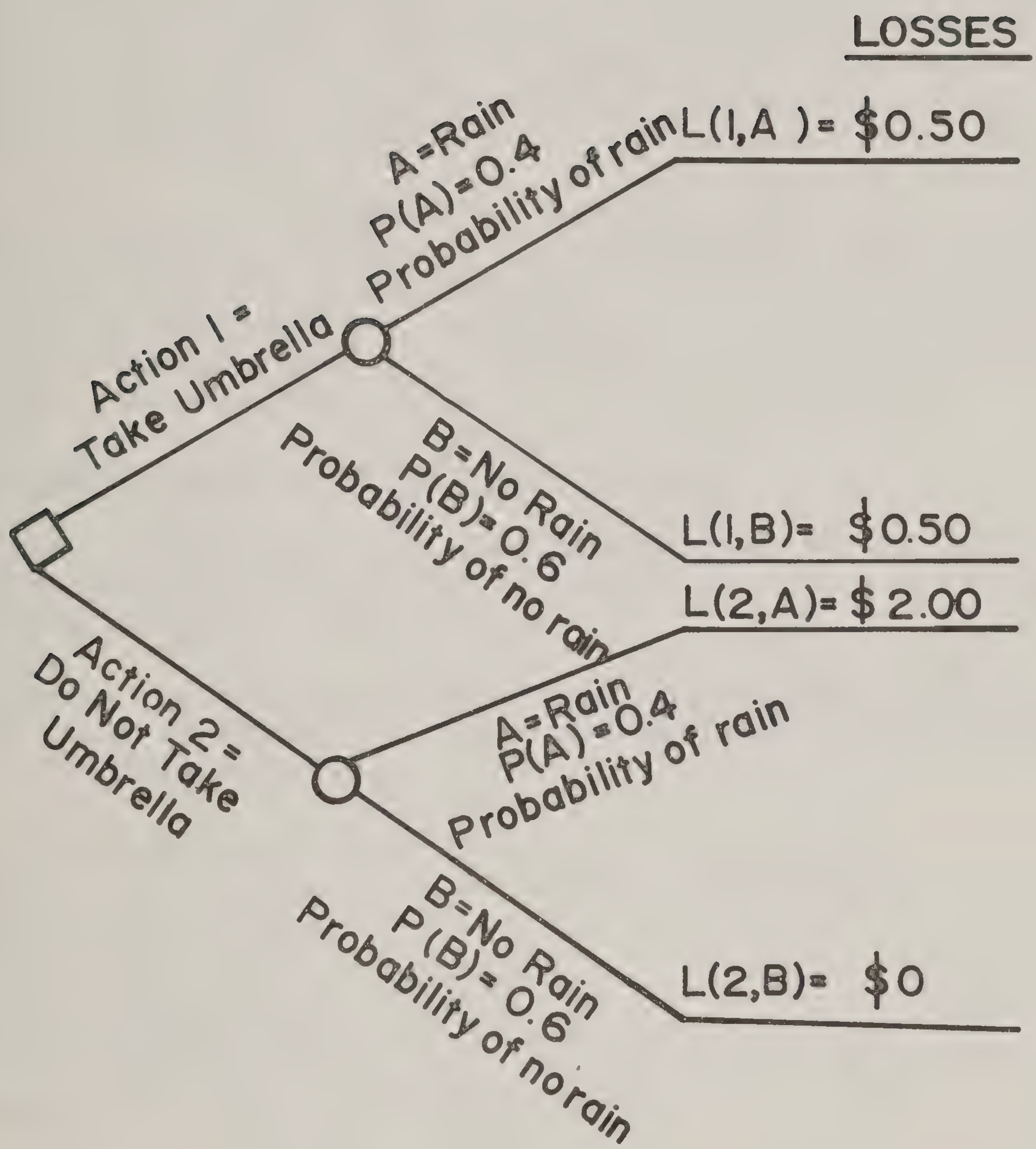
L(1,A) = Loss if umbrella is taken (action 1) and rain occurs (situation A)
= \$0.50 cloakroom fee

L(1,B) = Loss if umbrella is taken (action 1) and no rain occurs (situation B)
= \$0.50 cloakroom fee

L(2,A) = Loss if no umbrella is taken (action 2) and rain occurs (situation A)
= \$2.00 suit pressing charge

L(2,B) = Loss if no umbrella is taken (action 2) and no rain occurs (situation B)
= 0

The above information can be presented in a diagram termed a decision tree, which allows an orderly presentation of the action choices and the loss consequences due to future uncertain actual conditions.



☐ DECISION POINT

○ EVENT POINT

Here the gentleman, as the decision maker, stands at the trunk of the tree: the main branch system displays the actions which may be selected; and the secondary branch system displays the possible future conditions along with their corresponding probabilities and loss consequences. This decision tree presentation allows the decision maker to make an orderly visual analysis of any of the different combinations of actions and actual future conditions. For example, if we follow along the action 1 branch, we have a complete display of both the probability of outcome A, P(A) and the loss consequences L(1,A) if action 1 is taken and A occurs.

In actual engineering decision situations, where more than two actions and conditions may occur, the decision tree offers the best means of presenting all information to a group of decision makers. Each action, actual condition, probability, and loss consequence can be discussed individually and compared with the other information on separate tree branches.

(b) <u>Choice of Action based on Minimum Expected Loss</u>

In our simple example, the gentlemen will select the best action based on the criterion of minimum expected loss. Each action (J) has an expected, or average loss given by

$\bar{L}$ (J) = Sum of the Products of the losses (consequences) which may occur if action J is chosen multiplied by the probabilities of their occurrence.

$$\bar{L}(J) = L(J,A)P(A) + L(J,B)P(B)$$

For our particular problem

J = 1, Take umbrella, and
J = 2, Do not take umbrella,

the expected losses are:

$\bar{L}(1)$ = L(1,A)P(A) + L(1,B)P(B), if umbrella is taken
= (loss if rain)(probability of rain) + (loss if no rain)(probability of)
with unbrella   with umbrella   no rain
= (0.50)(0.4) + (0.50)(0.6)
= \$0.50

$\bar{L}(2)$ = L(2,A)P(A) + L(2,B)P(B), if umbrella is not taken
= (loss if rain)(probability of rain) + (loss if no rain)(probability of)
without umbrella   without umbrella   no rain
= (2.00)(0.40) + (0)(0.60)
= \$0.80

The action with minimum expected loss is action 1, take umbrella, and it is this action that the gentleman should accept.

The logic for this strategy of choosing that particular action with minimum expected loss is that if this minimum expected loss selection criteria is employed in a large number of action selection problems, the actual total loss will be less than if any other method of action selection were to be employed.

Clearly, the gentleman has no mental anguish in making decisions once that he has decided to let the minimum expected loss strategy dominate his life. All that he requires is the probabilities of actual future conditions P(A), P(B), and the values of loss due to actions and conditions L(1,A), L(2,A), L(1,B), L(2,B). 1/

A very interesting extension of the decision problem involves the selection of investigations which can reduce the uncertainty of future or unknown conditions. For example, is it worthwhile for the gentleman to invest $0.10 for a telephone call to the weather bureau, if their prediction of "rain" or "no rain" has a chance of 0.90 of being correct? Problems such as this are termed as "Pre-Posterior Analysis" since they involve selection of investigations previous to the "Posterior Analysis" of action selection, and the method of solution will be developed in Part II of this report.

Now, however, civil engineering examples will be presented in order to extend the minimum expected loss concept to a more important class of practical problems.

In all of the examples, the value or utility of actions and conditions will be expressed in terms of engineering estimates of cost or losses, since most engineers can easily understand how these values may be obtained. However, it should be understood that the minimum expected loss, or maximum expected gain concept can employ "utility values" which can represent social and environmental benefits on a rating scale which may differ from conventional monetary values.

A Dam Spillway Capacity Selection Problem

An earth dam is to be designed for a 100 year life. The designer must decide whether to select one of the following actions:

1 = construct 1000 cfs spillway capacity or
2 = construct 1500 cfs spillway capacity, at a cost of .5 million dollars more than the 1000 cfs spillway capacity dam.

1/ If the nature of the problem requires estimating the losses or probabilities of future conditions then careful estimates are essential

actual future conditions are:

A = Flood flow greater than 1000 cfs. (greater flow causes wash out of dam and complete destruction)

B = flood flow less than 1000 cfs.

C = flood flow greater than 1500 cfs.

D = flood flow less than 1500 cfs.

hydrological records provide the condition probabilities:

$P(A) = 0.2$

$P(B) = 0.8$

$P(C) = 0.1$

$P(D) = 0.9$

and economic analysis of construction and flood damage costs provide the losses:

$L(1,A)$ = 10 million dollars

$L(1,B) = 0$

$L(2,C)$ = 10.5 million dollars

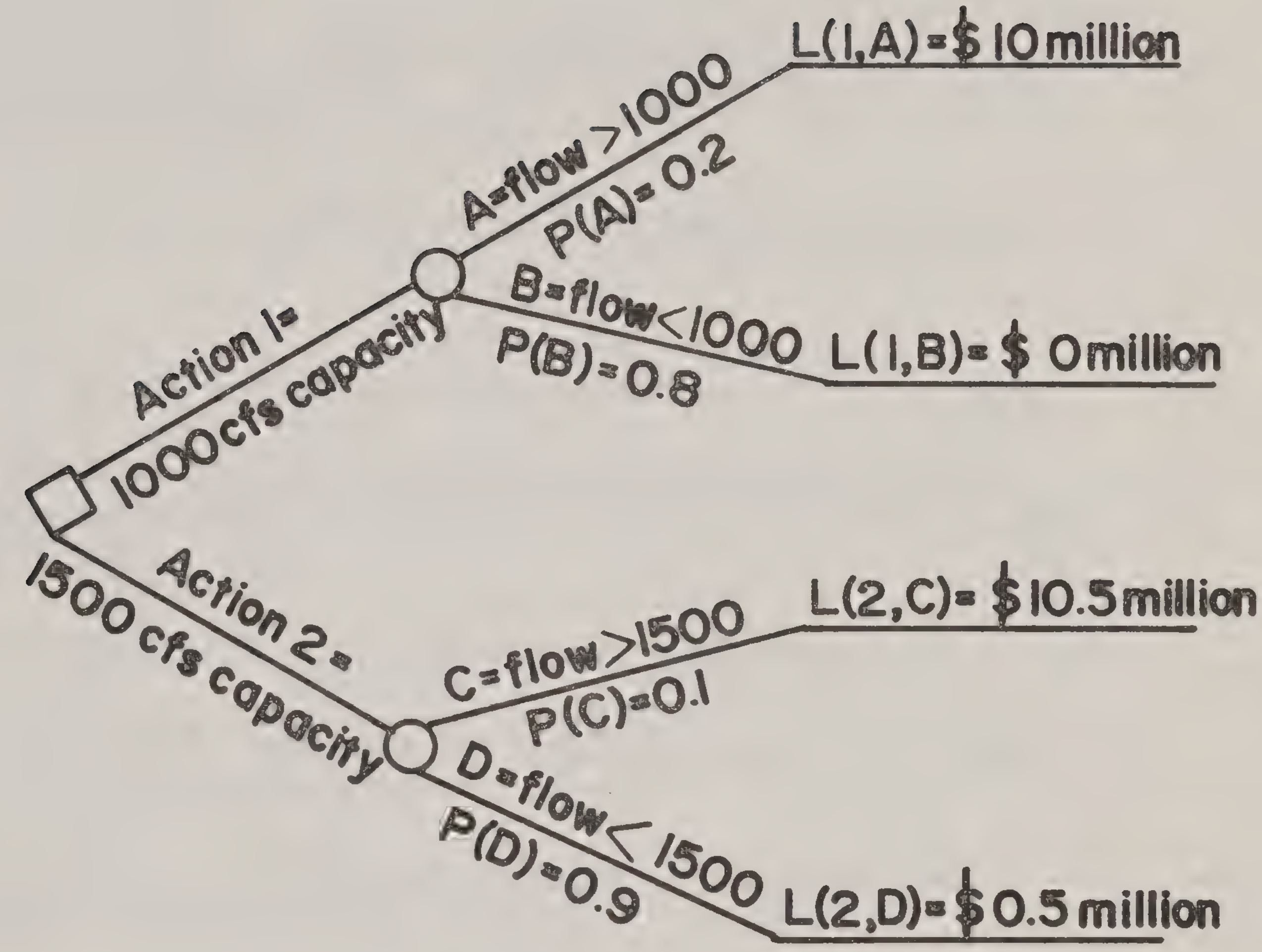
$L(2,D)$ = .5 million dollars

The expected losses are:

$\bar{L}(1) = 0.2\ (10) + 0.8\ (0) = 2$ million

$\bar{L}(2) = 0.1\ (10.5) + .09(0.5) = 1.50$ million

Here the minimum expected loss action is #2 = "construct 1500 cfs capacity."



Action 1 =
1000 cfs capacity
A = flow > 1000
P(A) = 0.2
L(1,A) = $ 10 million
B = flow < 1000
P(B) = 0.8
L(1,B) = $ 0 million
Action 2 =
1500 cfs capacity
C = flow > 1500
P(C) = 0.1
L(2,C) = $ 10.5 million
D = flow < 1500
P(D) = 0.9
L(2,D) = $ 0.5 million

(c) A General Formulation of the Decision or Action Selection Problem

In this section, it is desired to formalize the decision making procedure involving the selection of one particular action from a set of two or more alternative actions. The entire procedure will be developed with the use of the decision tree model.

Again, it is best to present the general symbols in terms of a simple example as follows: A structural engineer is involved in the preliminary design of a reinforced concrete building. His decision problem involves the selection of one of the following actions:

$\{a_j\}$ =

a_1 - Construct a concrete slab floor on two-way beams

a_2 - Construct a flat slab concrete floor without beams

The decision of whether to take action a_1 or a_2 depends on the performance (strength, deflection, cost) of the structural floor system. The possible future conditions are:

$\{S_i\}$ =

S_1 - Perfect Performance of two-way beam slab

S_2 - Adequate Performance (slight vibration under load) of two-way beam slab

S_3 - Inadequate Performance (excess deflections and cracking) of two-way beam slab

S_4 - Perfect Performance of Flat Slab

S_5 - Adequate Performance of Flat Slab

S_6 - Inadequate Performance of Flat Slab

Now, the action selection must be made under uncertainty concerning the actual future conditions $\{S_i\}$. The engineer must be able to obtain or assign probability values $P(S_i)$ for these future conditions. Based on experience with previously constructed floor systems we have

$P(S_i)$ =

$P(S_1) = 0.6$

$P(S_2) = 0.3$ Note $P(S_1) + P(S_2) + P(S_3) = 1$ for two-way beam slab

$P(S_3) = 0.1$

$P(S_4) = 0.5$

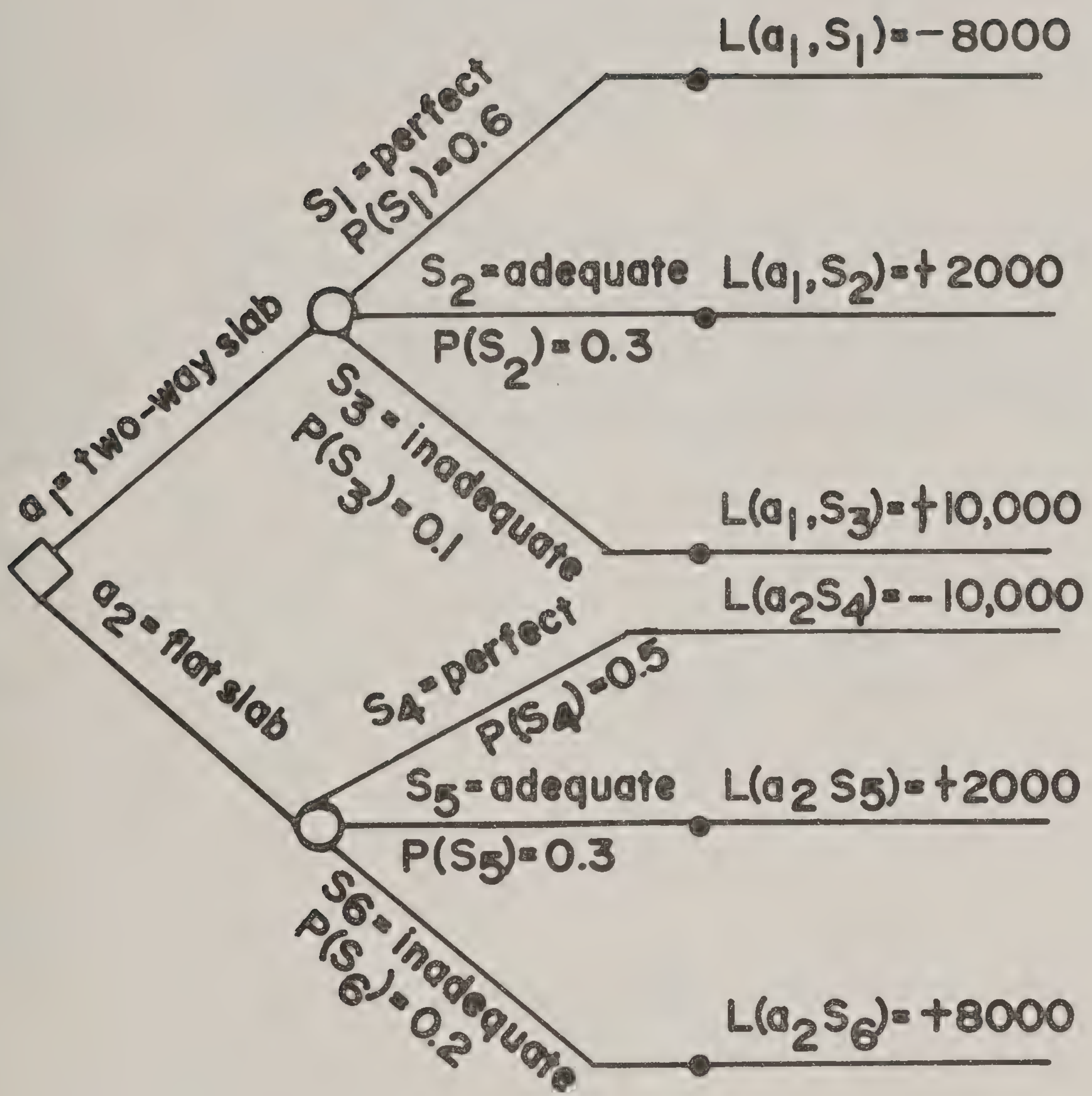
$P(S_5) = 0.3$ Note $P(S_4) + P(S_5) + P(S_6) = 1$ for flat slab

$P(S_6) = 0.2$

Loss is incurred by the engineer according to the action a_j he takes and the condition S_i which eventually occurs. For example, if the action is a_1, and S_3 occurs, the loss $L(a_1, S_3)$ is the consequence. The engineer must be able to estimate the loss values of each action-condition combinations.

Again experience with past construction projects provides the loss values for our example:

	$L(a_1,S_1)$	=	-\$8,000 (here a negative loss indicates a profit)
	$L(a_1,S_2)$	=	+2,000 (positive values are losses)
$L(a_j, S_i) =$	$L(a_1,S_3)$	=	+10,000
	$L(a_2,S_4)$	=	-10,000
	$L(a_2,S_5)$	=	+2,000
	$L(a_2,S_6)$	=	+8,000


a_1 = two-way slab
a_2 = flat slab
S_1 = perfect
$P(S_1) = 0.6$
S_2 = adequate
$P(S_2) = 0.3$
S_3 = inadequate
$P(S_3) = 0.1$
S_4 = perfect
$P(S_4) = 0.5$
S_5 = adequate
$P(S_5) = 0.3$
S_6 = inadequate
$P(S_6) = 0.2$
$L(a_1, S_1) = -8000$
$L(a_1, S_2) = +2000$
$L(a_1, S_3) = +10{,}000$
$L(a_2 S_4) = -10{,}000$
$L(a_2 S_5) = +2000$
$L(a_2 S_6) = +8000$

With this tree diagram we can follow through the entire possible condition and loss consequence sequence for any selected action or decision.

A decision is thus defined as the taking of a particular action a_j from the set of all possible actions, $\{a_j\}$. The results or outcome of the action a_j is the condition S_i which is random and has the probability of occurrence $P(S_i)$. After taking a_j, the S_i occurs and loss $L(a_j, S_i)$ is incurred. The selection of the best action a_j proceeds as follows.

With the loss values $L(a_j, S_i)$ and probabilities $P(S_i)$, it is possible to compute the expected value of each action a_j. The engineer will then take the action which has the minimum expected loss. The expected loss of an action a_j is found by the multiplication of the individual probabilities of condition $P(S_i)$ by the corresponding loss $L(a_j, S_i)$ for the action and the condition and then summing these products.

$$\overline{L}(a_j) = \sum_{\text{for all } S_i} L(a_j, S_i) \cdot P(S_i)$$

(where $\sum_{\text{for all } S_i}$ means the sum of all products for each S_i)

For the floor slab example,

$$\overline{L}(a_1) = \Sigma\, L(a_1, S_i) \cdot P(S_i)$$

$$= L(a_1, S_1) \cdot P(S_1) + L(a_1, S_2) \cdot P(S_2) + L(a_1, S_3) \cdot P(S_3)$$

$$= (-8) \cdot (0.6) + (2) \cdot (0.3) + (10) \cdot (0.1) = -3.2$$

(where the (-) indicates an expected profit)

$$\overline{L}(a_2) = \Sigma\, L(a_2, S_i) \cdot P(S_i)$$

$$= L(a_2, S_4) \cdot P(S_4) + L(a_2, S_5) \cdot P(S_5) + L(a_2, S_6) \cdot P(S_6)$$

$$= (-10) \cdot (0.5) + (2) \cdot (0.3) + (8) \cdot (0.2) = -2.8$$

Here the minimum expected loss or maximum expected profit action is a_1.

Evaluation of the Probabilities $P(S_i)$ of Future or Unknown Conditions or Situations

In the previous examples the values for the probabilities $P(S_i)$ were found by:

(1) Weather Department Experienced Judgement of the chances of rain, for the umbrella decision.

(2) An empirical plot of a historical record flood flow data on probability paper for the dam spillway decision.

(3) Past performance experience of the engineer with concrete floor systems allowed the assignment of the chances of levels of performance by judgement, in the slab decision problem.

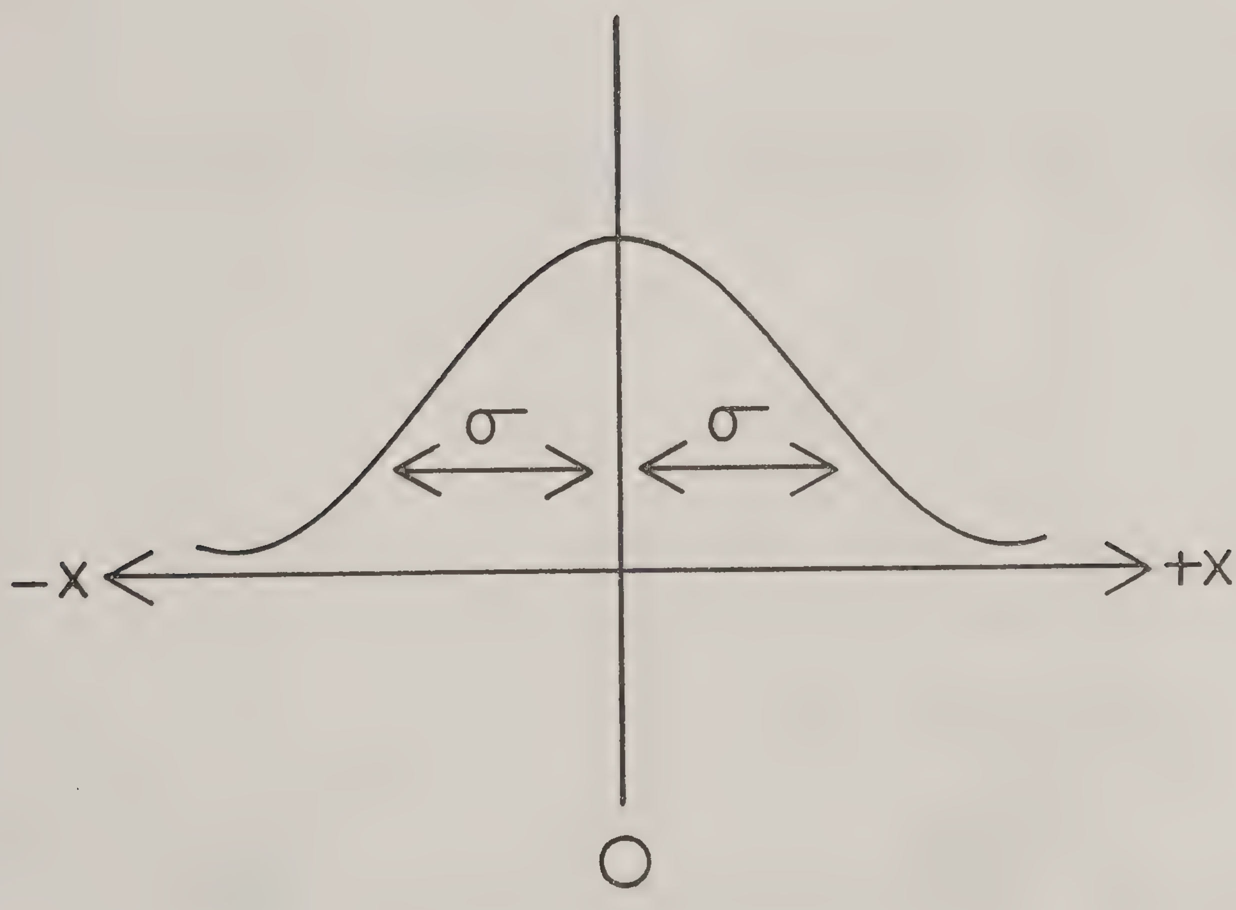
In most engineering applications the probability values will be assigned, as in this latter case, by the judgement of an expert with substantial experience concerning the future unknown condition.

However, in this next example we will see that occasionally, probabilities can be computed in terms of accepted technical definitions, such as: "permissible error must not be exceeded with 90% reliability."

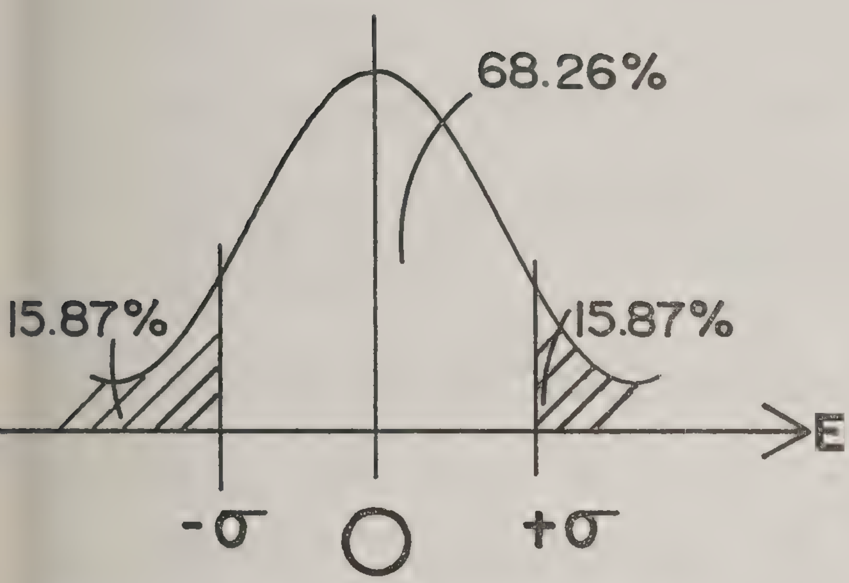
An Example from Surveying

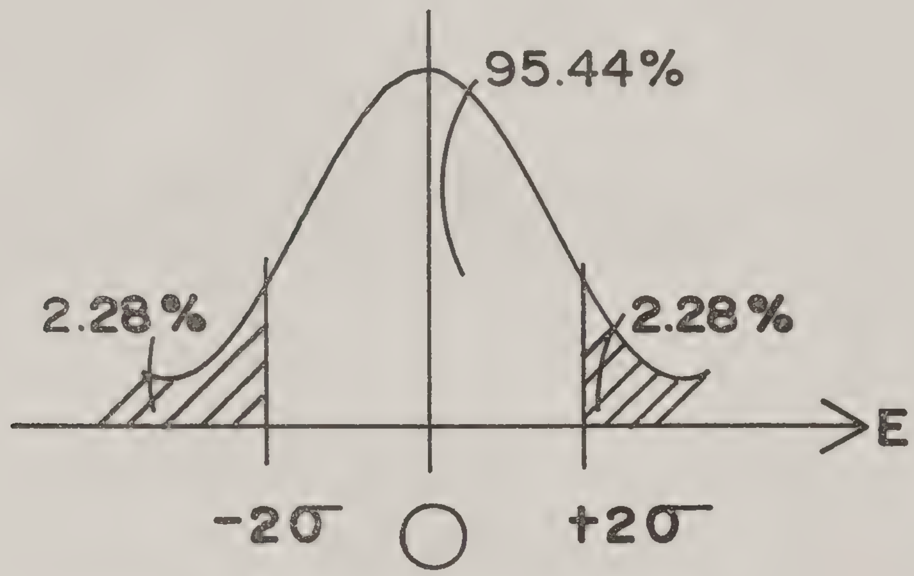
A civil engineer is engaged in a surveying job concerning the running of a one mile length of levels. His decision problem is to select the use of either first order work or second order work. Specifically, the engineer's client requires that the one-mile length of levels be run with an error of closure less or equal to 0.015 feet. The level circuit must be repeated if a greater error occurs. A time limit does not allow more than one repeat to be made.

a) Probability of Error: The final error of closure X is a random variable. This error is made up of the sums of the individual errors made at each rod reading in the circuit. A random variable such as X, which is a sum of a large number of random rod reading errors, has a "Normal" or "bell-shaped" probability distribution.



This curve is centered on the zero algebraic average value of error, representing the fact that error can be either positive or negative with equal likelihood. The spread of the curve is represented by the standard error σ, and probabilities of errors within given limits are provided by areas under the curve. For example:

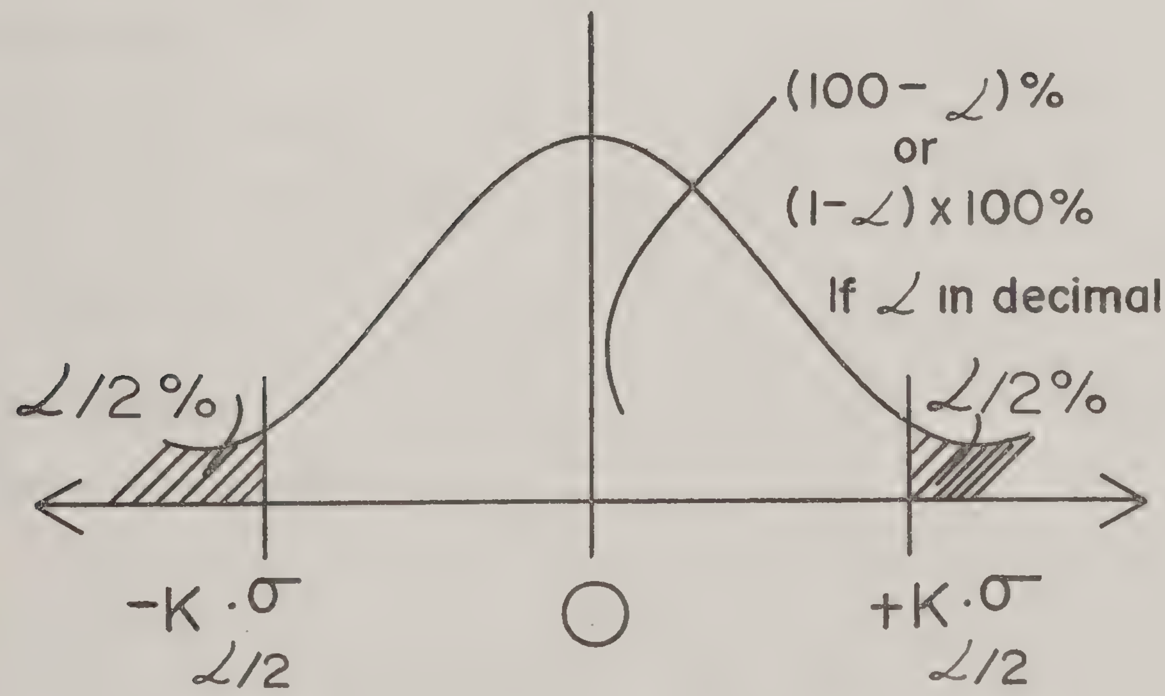




Prob $[-\sigma \leq E \leq +\sigma]$ = 68.26%

Prob $[-2\sigma \leq E \leq 2\sigma]$ = 95.44%

In fact, Normal Probability Tables (see Reference (2), page 555) can be used to find the area under any specified multiple of the standard error σ.



Prob $[-K_{\alpha/2}\sigma \leq E \leq +K_{\alpha/2}\sigma]$ = $(1 - \alpha)\%$

Thus with Normal Tables and a given value of σ we can find the error limits $\pm K_{\alpha/2}\sigma$ such that the probability of remaining within these limits is a given value of $(1 - \alpha)\%$. Or, for a given value of $(1 - \alpha)\%$ value we can evaluate the standard error σ. This is applicable to our particular problem:

First and second order level work are specified procedures (see Reference (3) pages 9 and 12) such that for a 1 mile length of levels,

<u>First Order Permissible Error</u> = E_1

$E_1 = 0.012$ ft.

Prob $[-0.012 \leq E \leq +0.012] = 90\%$

This means that from experience, the error has been found to remain within ± 0.012 ft. within 90% of the time when First Order work procedures are employed.

<u>Second Order Permissible Error</u> = E_2

$E_2 = 0.025$ ft.

Prob $[-0.025 \leq E \leq +0.025] = 90\%$

The above specifications for the $(1 - \alpha)\% = 90\%$ reliability of permissible error allows the determination of the standard errors σ_1 and σ_2 for first and second order work respectively. Normal Probability Tables give the multiple $K_{\alpha/2} = K_{5\%} = 1.64$ such that

$$\text{Prob } [-1.64\sigma \leq E \leq +1.64\sigma] = 90\%$$

and therefore for:

<u>First Order Work</u>

$$1.64\sigma_1 = E_1 = 0.012$$

giving the first order standard error,

$$\sigma_1 = \frac{0.012}{1.64} = 0.0073$$

<u>Second Order Work</u>

$$1.64\sigma_2 = E_2 = 0.025$$

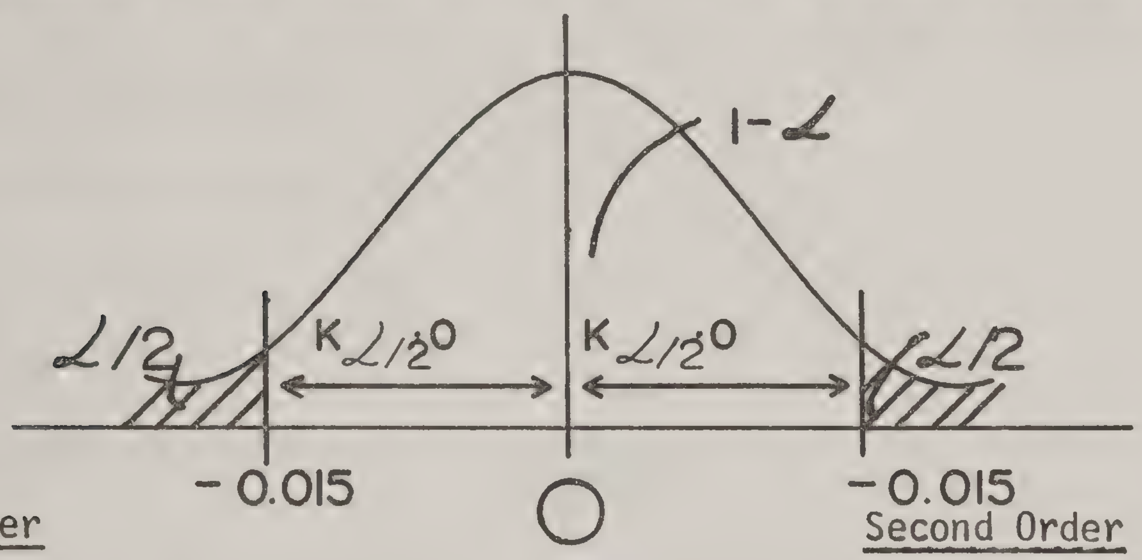
giving the second order standard error,

$$\sigma_2 = \frac{0.025}{1.64} = 0.0152$$

With these σ values the probabilities

$$P[-0.015 \leq E \leq +0.015]$$

can be evaluated from Normal Tables for first and second order work as follows,



First Order

$$K_{\alpha/2} = \frac{0.015}{\sigma_1} = \frac{0.015}{0.0073} = 2.06$$

$\alpha/2 - .02 = 2\%$ (From Tables)

$\alpha = -.04$

$1-\alpha = 0.96 = 96\%$

Second Order

$$K_{\alpha/2} = \frac{0.015}{\sigma_2} = \frac{0.015}{0.0152} = 0.99$$

$\alpha/2 = 0.161 - 16.1\%$ (From Tables)

$\alpha = 0.322$

$1-\alpha = 0.678 - 67.8\%$

Defining the actual future consitions $\{S_i\}$, ($|E|$ is absolute value of E)

$S_1 = |E| \leq 0.015$, with first order work,

$S_3 = |E| \leq 0.015$, with second order work

$S_2 = |E| > 0.015$, with first order work,

$S_4 = |E| > 0.015$, with second order work

Defining the Actions $\{a_j\}$

a_1 = First Order Work

a_2 = Second Order Work

The Probabilities are as found above,

$P(S_1) = 0.96$ $\quad$ $P(S_3) = 0.678$

$P(S_2) = 0.04$ $\quad$ $P(S_4) = 0.322$

b) The Loss Calculations:

Both orders require 5 man crew:

2 men at $5/hr	=	10
3 men at $5/hr	=	9
Wild Level at $1/hr	=	1
		$20/hr

First order requires shorter shot lengths, i.e., more time

time = 8 hrs at 20 = $160

Second order

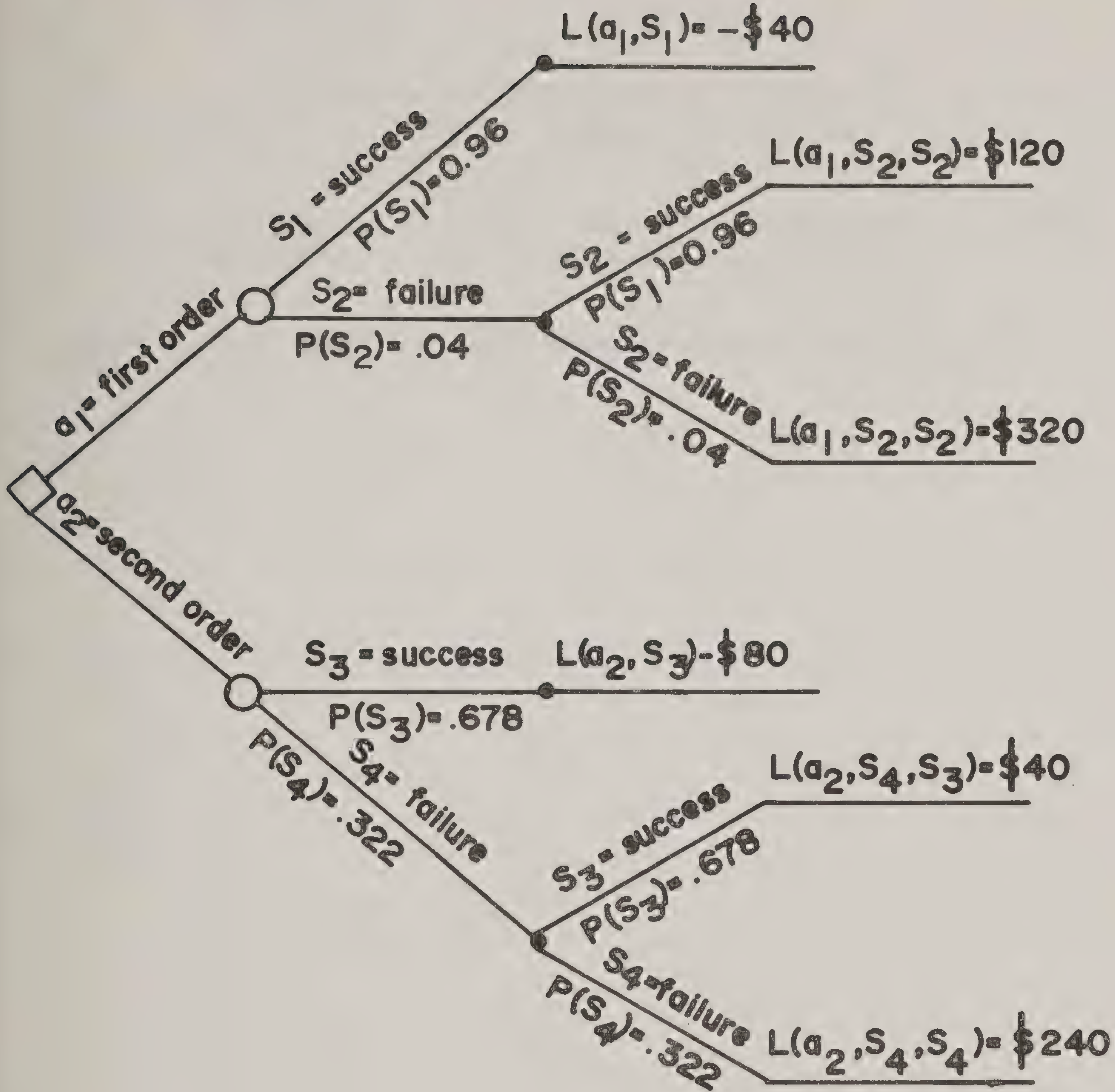
time = 6 hrs at 20 = $120

Engineer appraises this job at $200

Note: It is assumed that once the engineer begins with a given order then he will use that same order again if the first outcome is unsuccessful ($E > .015$)

The actions, outcomes and losses are as follows:

Action	Outcome	2nd outcome if 1st unsuccessful	Loss
a_1	S_1(success)	--	$L(a_1,S_1) = -200 + 160 = -\$40$ (a profit)
a_1	S_2(failure)	S_1(success)	$L(a_1,S_2,S_1) = -200 + 160 + 160 = \$120$ (a loss)
a_1	S_2(failure)	S_2(failure)	$L(a_1,S_2,S_2) = 160 + 160 = \$320$
a_2	S_3(success)	--	$L(a_2,S_3) = -200 + 120 = -\$80$
a_2	S_4(failure)	S_3(success)	$L(a_2,S_4,S_3) = -200 + 120 + 120 = \$40$
a_2	S_4(failure)	S_4(failure)	$L(a_2,S_4,S_4) = 120 + 120 = \$240$


L(a1,S1)= −$40
S1 = success
P(S1)=0.96
L(a1,S2,S2)=$120
S2 = success
P(S1)=0.96
S2= failure
P(S2)= .04
a1= first order
S2= failure
P(S2)= .04
L(a1,S2,S2)=$320
a2= second order
S3 = success
L(a2,S3)-$80
P(S3)= .678
P(S4)= .322
S4= failure
L(a2,S4,S3)=$40
S3= success
P(S3)= .678
S4=failure
P(S4)= .322
L(a2,S4,S4)= $240

d) The Expected Loss:

$$\bar{L}(a_1) = \underset{\text{(success)}}{(.96)(-40)} + \underset{\text{(failure)}}{(.04)}\ [\underset{\text{(success)}}{(.96)(120)} + \underset{\text{(failure)}}{(.04)(320)}] = -\$33.28 \text{ (an expected profit)}$$

$$\bar{L}(a_2) = \underset{\text{(success)}}{(.678)(-80)} + \underset{\text{(failure)}}{(.322)}[\underset{\text{(success)}}{(.678)(40)} + \underset{\text{(failure)}}{(.322)(240)}] = -\$20.62$$

Choose a_1, First Order Work, with minimum expected loss or maximum expected profit.

II. DECISIONS CONCERNING THE SELECTION OF AN INVESTIGATION (PRE-POSTERIOR ANALYSIS)

We return now to the umbrella problem given in part I, by asking whether it is worthwhile to invest \$0.10 for a telephone call to the weather bureau, if the probability of their predicting rain when indeed it will rain and the probability of their predicting no rain when indeed it will not rain are both .90.

Before solving this problem we will introduce some needed notation (1)-(4) and formulas (5)-(7).

(1) $P(AB)$ = The probability that both A and B will occur.

(2) $P(\bar{A})$ = The probability that A will not occur.

(3) $P(A\bar{B})$ = The probability that A will occur and B will not occur.

(4) $P(B|A)$ = The probability (conditional probability) that B will occur given that A has occurred.

$P(B|A)$ may be found by determining $P(BA)$ and $P(A)$ and then dividing the former by the latter as shown in formula (5)

$$(5)\quad P(B|A) = \frac{P(BA)}{P(A)}$$

Given $P'(A)$, $P(B)$ and the conditional probability $P(B|A)$ the "reverse" conditional probability $P''(A|B)$ may be found by formula (6) below: (See reference (1) p. 119)

$$(6)\quad P''(A|B) = \frac{P(B|A)\,P'(A)}{P(B)}$$

, where superscripts (', ") are used to illustrate the following differences in these probabilities. $P'(A)$ (usually known) is the original (prior) probability that A will occur, whereas $P''(A|B)$ is the revised (posterior) conditional probability that A will occur given that more information has been gathered and it has been found that B has occurred.

Finally formula (7) below states that the probability that A will occur is the sum of the probability that A and B will occur and the probability that A and not B will occur.

$$(7)\quad P(A) = P(AB) + P(A\bar{B})$$

We have seen that if no call has been made to the weather bureau then the best course of action, based upon the minimum expected loss (\$0.50), is "take umbrella". We now determine what the expected loss will be if a call is made to the weather bureau followed by the best further action after their report. We then choose to call the weather bureau if this expected loss is less then \$0.50.

Below are tabulated the given probabilities.

Prior Probabilities

$P'(R) = .4$; the prior probability of rain

$P'(\overline{R}) = .6$; the prior probability of no rain

Conditional probabilities we wish to "reverse" to obtain P"(rain|report rain) etc.

$P(RR|R) = .9$; the probability of reporting it will rain when it will.

$P(\overline{RR}|R) = .1$; the probability of reporting it will not rain when it will.

$P(RR|\overline{R}) = .1$; the probability of reporting it will rain when it will not.

$P(\overline{RR}|\overline{R}) = .9$; the probability of reporting it will not rain when it will not.

Below we determine the unknown probabilities.

The following two probabilities needed for determining the "reversed" posterior probabilities are:

$P(RR)$; the probability of reporting it will rain and

$P(\overline{RR})$; the probability of reporting it will not rain.

They are as follows:

$$\begin{aligned}P(RR) &= P(RR \cdot R) + P(RR \cdot \overline{R}) \\ &= P(RR|R)P'(R) + P(RR|\overline{R})P'(\overline{R}) \\ &= (.9)(.4) + (.1)(.6) \\ &= .42\end{aligned}$$

Hence $P(\overline{RR}) = 1-P(RR) = .58$

We now determine the posterior probabilities of rain or no rain given the weather report as follows:

$P''(R|RR)$ = the probability of rain given a report of rain

$$= \frac{P(RR|R)P'(R)}{P(RR)} = \frac{(.9)(.4)}{.42} = .857$$

Hence:

$P''(\overline{R}|RR)$ = the probability of no rain given a report of rain

$= 1 - P''(R|RR) = .143$

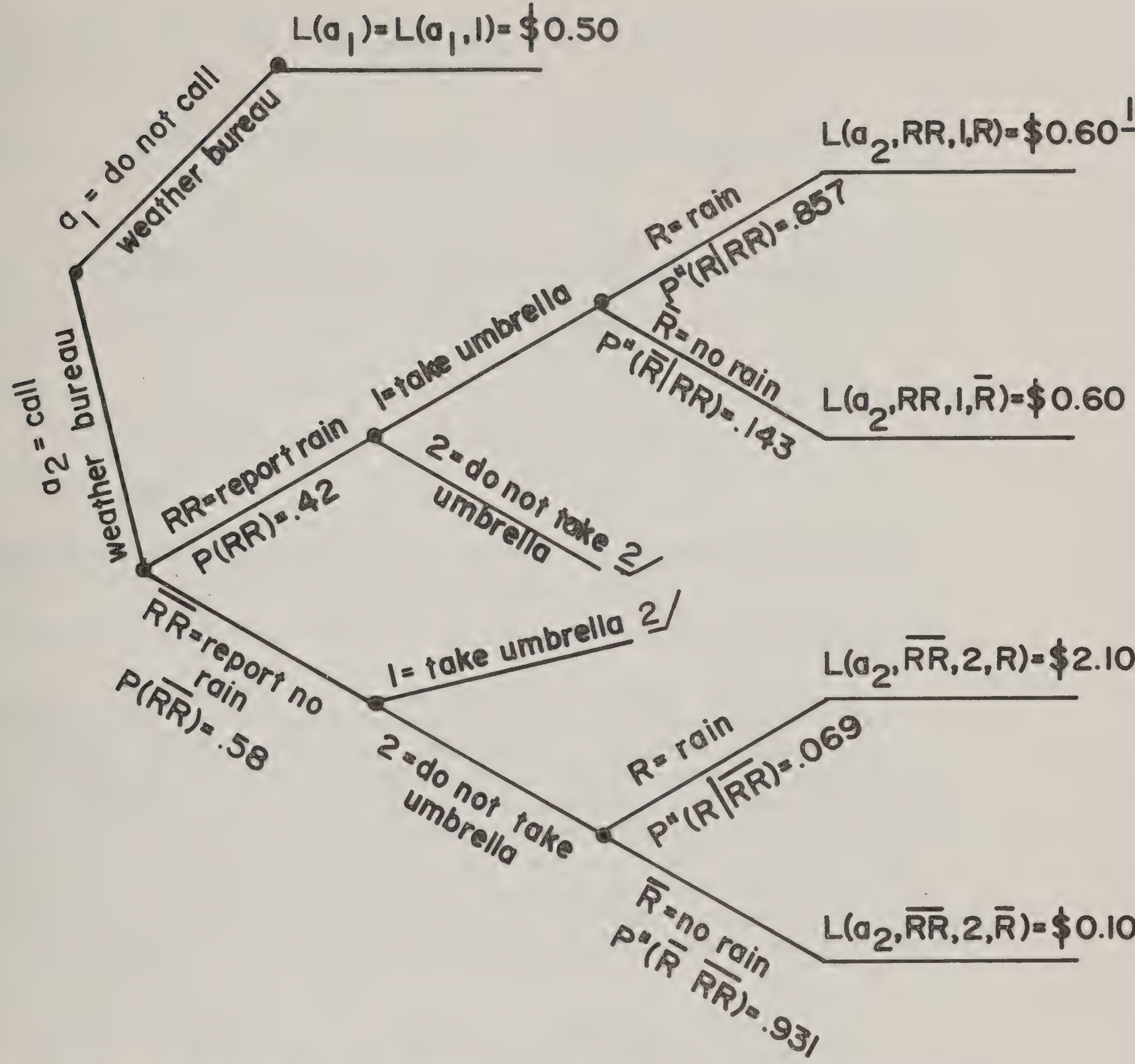
$P''(R|\overline{RR})$ = the probability of rain given a report of no rain

$$= \frac{P(\overline{RR}|R)P'(R)}{P(\overline{RR})} = \frac{(.1)(.4)}{(.58)} = .069$$

Hence:

$P''(\overline{R}|\overline{RR})$ = the probability of no rain given a report of no rain

$= 1 - P''(R|\overline{RR}) = .931$



1/ $L(A_2, RR, 1, R)$ is the loss resulting from calling the weather bureau and obtaining a report of rain and taking the unbrella when it indeed rains.

2/ Note: These two actions are those ruled out by using the principle of least expected loss at that action juncture. You may compute these two expected losses as an exercise. The significance of this is that for each possible result of every investigation candidate, the best choice of further action is both (1) necessary for selecting the best method of investigation and (2) available for immediate use after the results of the best investigation. This final step is called terminal analysis and is discussed later in this section.

The expected loss $\bar{L}(a_2)$ resulting from investigation a_2 "Call the weather bureau" is:

$$\bar{L}(a_2) = P(RR)[P''(R|RR)L(a_2,RR,1,R) + P''(\bar{R}|RR)L(a_2,RR,1,\bar{R})]$$

$$+$$

$$P(\overline{RR})[P''(R|\overline{RR})L(a_2,\overline{RR},2,R) + P''(\bar{R}|\overline{RR})L(a_2,\overline{RR},2,\bar{R})]$$

$$= (.42)[(.857 \times \$0.60 + (.143) \times \$0.60)]$$

$$+$$

$$(.58)[(.069) \times \$2.10 + (.931) \times \$0.10]$$

$$= \$0.39$$

Hence we conclude that it is best to call the weather bureau based upon the principle of least expected cost.

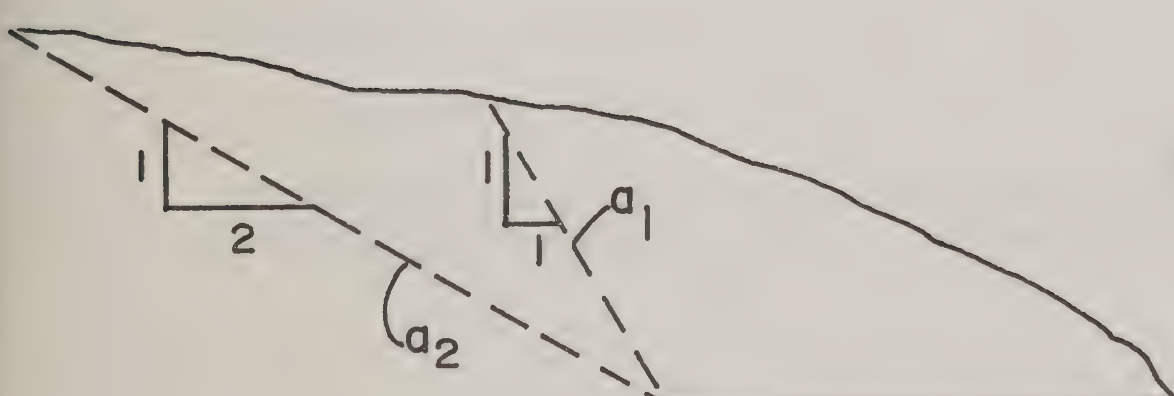
AN EXAMPLE INVOLVING CHOICE OF A GEOLOGICAL SITE INVESTIGATION

In the preparation of a road construction contract proposal, the engineer must decide which type of cut slope to specify for a 1/4 mile length of hill-side cut. The type of cut slope depends upon the unknown orientation condition of geological strata at the cut; where differing orientations of strata can cause different rock slide problems on the cut slope. The engineer can obtain information about the unknown strata by ordering one of two levels of geological site investigations, and therefore the pre-posterior decision analysis is to select the optimum type of site investigation. Detailed descriptions of the actions, orientation conditions, investigations, results, losses, and probabilities are given as follows.

A. Set of Actions $\{a_j\} = \{a_1, a_2\}$

a_1 = 1 : 1 Cut Slope (to be called for in the contract proposal)

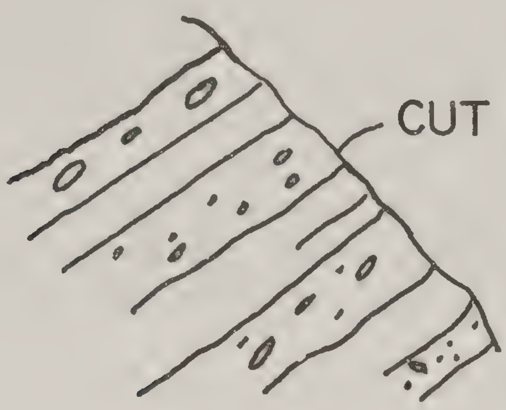
a_2 = 2 : 1 Cut Slope



Note: Cut volumes are substantially different for each type of specified slope.

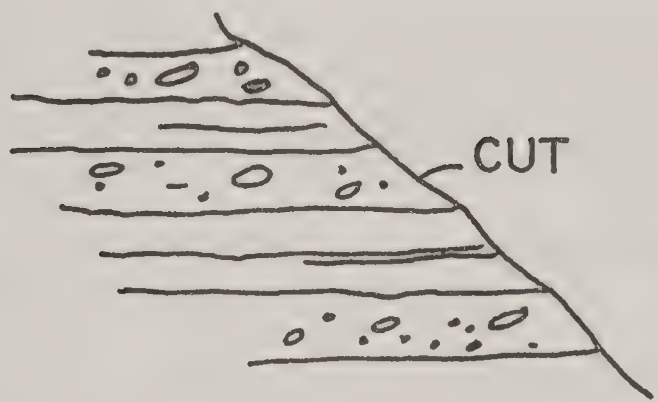
B. Orientation Conditions $\{S_i\} = \{S_1, S_2\}$

S_1 = Orthogonal (to cut) Strata



Note: Can use a_1 = 1:1 slope for good slide control

S_2 = Horizontal Strata



Note: Can use a_2 = 2:1 slope for good slide control

C. List of Possible Site Investigations $\{e_k\} = \{e_1, e_2\}$

e_1 = Hire geological consultant with manual sub-surface exploration by shovels and augers.

e_2 = Hire geological consultant with drilling rig for deep cores.

D. Set of Possible Results of Investigations $\{z_\ell\} = \{z_1, z_2\}$

z_1 = Indication that S_1 exists

z_2 = Indication that S_2 exists

Note: An indication does not definitely establish that a given condition exists. It may, however, increase or decrease the probability of the presence of a given orientation condition S_i.

E. Probability Assignments

1. Prior Probabilities for the Unknown Orientation Conditions:

In order to select one of the two types of investigations e_k, the engineer must have some idea of the probability of the existence of

each unknown strata condition S_i. Now, since the engineer assigns a number value $P'(S_i)$ for these probabilities prior to the investigation, they are termed as "Prior Probabilities of Conditions" = $P'(S_i)$. Here it is necessary for the engineer to employ all of his technical knowledge and experience to best evaluate the prior chances $P'(S_i)$ before having any knowledge of investigation results. Of course, there will be cases where the engineer or decision maker will be totally ignorant of the conditions. This state of total ignorance provides a base line for probability assignment; for example if the engineer has no knowledge about orientation conditions S_1 or S_2, he should assign equal probabilities, totaling up to unity,

$$P'(S_1) = 1/2 \ , \ P'(S_2) = 1/2$$

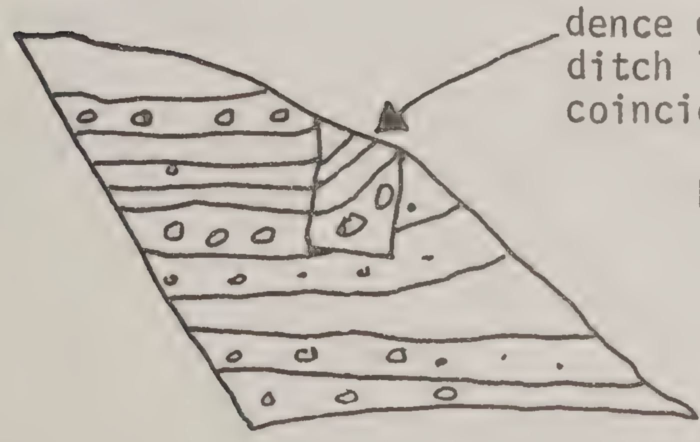
That is, total ignorance may be expressed by equal chances for all unknown conditions.

In our problem however, the engineer has seen the character of the terrain in his past travels and has a slight preference for the horizontal orientation S_2, such that his prior probability assignments are $P'(S_1) = 0.4$, and $P'(S_2) = 0.6$.

2. Conditional Probabilities of Investigation Results for a Given Investigation and Actual Condition:

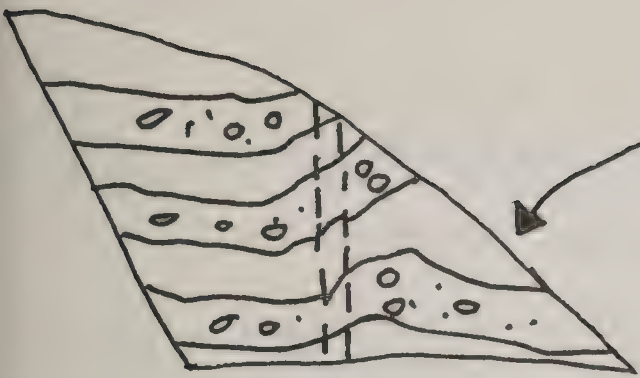
In many types of engineering investigations, technical knowledge and experience will allow the assignment of a Conditional Probability $P(z_\ell|S_i, e_k)$ of obtaining a certain result z_ℓ from a given investigation e_k when an actual known condition S_2 exists. The term "conditional probability" implies that the probability is valid only for the condition that S_i is the true existing situation.

The given condition S_i allows us either to draw physically, or describe mathematically, the true given condition S_i, and then estimate or calculate the chance of obtaining a given result z_ℓ with investigation e_k. For example, the geologist is able to draw the actual S_2 condition of horizontal strata on the roadway cut cross-sections, then show the types of his investigations e_1 or e_2; and then estimate the chances of getting z_1 or z_2 from each of these investigations; where natural variability of the horizontal strata might yield erroneous results due to just "bad luck" at the investigation sites.



Hand Dug Ditch Investigation e_1 may yield erroneous z_1 indication here due to coincidence of upward strata variation at the ditch location. Chance of this "bad luck" coincidence is estimated as

$$P(z_1|S_2, e_1) = 0.3$$



Drill Hole Boring Investigation e_2 may yield erroneous z_1 indication here. Chance of this "bad luck" coincidence of a boring with an upward variability of strata is estimated by the geologist as

$$P(z_1|S_2, e_2) = 0.1$$

The geologist is able to assign this estimated probability value because he has the knowledge of the frequency of "upward strata irregularities" that may occur when the true strata is horizontal S_2.

For the assignment of $P(z_\ell|S_i, e_k)$, in general, the investigator asks himself the following question: "What is the probability that my investigation e_k will yield a result z_ℓ when the true condition is S_i?" or "With what reliability will my investigation e_k detect the condition S_i when it truly exists?" The fact that these conditional probabilities can be assigned by an experienced expert, with access to past records and technical knowledge relating to the given actual conditions and investigations, is a most important input phase of the computations leading to the pre-posterior selection of an optimum investigation. This is because the conditional probabilities $P(z_\ell| S_i, e_k)$ represent a numerical measure of the investigator's opinion concerning the accuracy or quality of the results of a given proposed investigation e_k. For example, if the experience and knowledge concerning a certain investigation e_k is such that the investigator feels very confident that it will furnish a reliable indication z_i of the condition S_i when S_i really exists, then he may express this confidence numerically by assigning a high value for $P(z_i|S_i, e_k)$ such as 0.9. In the analysis procedure, this high $P(z_i|S_i, e_k)$ value will favor the actual selection of e_k as the optimum investigation, if its cost is not too high. The complete set of Conditional Probabilities are given below:

	method of investigation e_1		method of investigation e_2	
$P(z_\ell\|S_i, e_k)$	z_1	z_2	z_1	z_2 ← indicated condition
$P(z_\ell\|S_1, e_k)$	.7	.3	.9	.1
$P(z_\ell\|S_2, e_k)$	.3	.7	.1	.9
actual condition ↑				

For example: the geologist knows by experience that there is a 0.7 chance of z_1 when investigation e_1 is made and S_1 is the actual condition.

F. Losses or Costs $L(e_k, z_\ell, a_j, S_i)$ for each Combination of Investigation, Result, Action, and Actual Condition.

All costs are expressed in terms of Equivalent Annual Cost in Dollars.

Construction Costs: These vary for each action a_j because of the different earthwork cut volumes for each type of specified slope. These costs can therefore be estimated from known volumes and unit costs.

Construction Cost for a_j = $CC(a_j)$

$CC(a_1)$ = 0, here the cost for this minimum necessary action is used as the zero base.

$CC(a_2)$ = 200

Maintenance Costs: These vary for each combination of specified slope a_j, and actual orientation condition S_i. Here, particularly, past available maintenance records for each of the above combinations, can provide cost data.

Maintenance Cost for a_j, with S_i = $MC(a_j, S_i)$

$MC(a_1, S_1) = 0$; $MC(a_2, S_1) = 0$

Because each slope can adequately control slides with the S_1 orientation condition.

$MC(a_1, S_2) = 400$; $MC(a_2, S_2) = 0$;

Because a_1 = 1:1 slope does not control slides with the S_2 orientation condition.

Investigation Costs: These vary with the manpower, equipment and time required for each type of investigation e_k. Here it is assumed that the cost of any investigation e_k does not vary with the result or indication z_ℓ detected by the investigation.

Investigation Cost for investigation e_k with result z_ℓ = $IC(e_k, z_\ell)$

$$IC(e_1, z_\ell) = 100, \quad IC(e_2, z_\ell) = 200$$

All are valid for either result z_1, z_2.

Total Cost or Losses: The total loss for a given investigation e_k, result z_ℓ, action a_j, and condition S_i is

$$L(e_k, z_\ell, a_j, S_i) = IC(e_k, z_\ell) + MC(a_j, S_i) + CC(a_j).$$

This total cost or loss is given in the following table,

$L(e_k, z_\ell, a_j, S_i)$	a_1	a_2
$L(e_1, z_\ell, a_j, S_1)$	100	300
$L(e_1, z_\ell, a_j, S_2)$	500	300
$L(e_2, z_\ell, a_j, S_1)$	200	400
$L(e_2, z_\ell, a_j, S_2)$	600	400

The above costs do not vary with investigation results z_1, z_2.

A Flow Chart Guide for the Reader

The following Flow Chart is provided to serve as a general guide to the reader, during the simple, but lengthy, calculations involved in this numerical example. Definitions of all symbols will be given directly in the example.

It should be recalled that the basic objective is to decide which type of investigation e_k, should be made in order to improve the unknown condition probability values from the rough prior (prior to the investigation e_k) values of $P'(S_i)$ to refined or sharper posterior (posterior to having conducted e_k with a result of z_ℓ) values of $P''(S_i|z_\ell, e_k)$. For example, the engineer wants to know if it is worthwhile to spend $200 on the boring investigations e_2, which has the reliability $P(z_1|S_1, e_2) = 0.9$ of providing an indication z_1 of upward strata when upward strata S_1 is the true condition; such that the prior probability of upward strata $P'(S_1) = 0.4$ would be changed to the sharper posterior value (to be calculated in the example) of $P''(S_1|z_1, e_2) = 0.86$.

FLOW CHART

FOR PRE-POSTERIOR ANALYSIS

Given the Input Quantities

$P'(S_i)$

$P(z_\ell | S_i, e_k)$

$L(e_k, z_\ell, a_j, S_i)$

Compute Probabilities

$$P(z_\ell | e_k) = \sum_{\substack{\text{sum} \\ \text{for all } S_i}} P(z_\ell | S_i, e_k) \cdot P'(S_i)$$

$$P''(S_i | z_\ell, e_k) = \frac{P(z_\ell | S_i, e_k) \cdot P(S_i)}{P(z_\ell | e_k)}$$

Compute Expected Losses for Actions

$$\bar{L}(e_k, z_\ell, a_j) = \sum_{\substack{\text{sum} \\ \text{for all } S_i}} L(e_k, z_\ell, a_j, S_i) \cdot P''(S_i | z_\ell, e_k)$$

then select the minimum value

$$\bar{L}(e_k, z_\ell, a^*) = \text{minimum of } \bar{L}(e_k, z_\ell, a_j)$$

Compute Expected Loss or Cost of Investigation

$$\bar{L}(e_k) = \sum_{\substack{\text{sum} \\ \text{for all } z_\ell}} \bar{L}(e_k, z_\ell, a^*) \, P(z_\ell | e_k)$$

Then select the minimum value

$$\bar{L}(e^*) = \text{minimum of } \bar{L}(e_k)$$

The optimum or minimum expected loss investigation is that e_k which has $\bar{L}(e^*)$

G. Calculation of Further Required Probabilities

In order to proceed with the selection of the optimum, "minimum average loss", investigation, it is necessary to calculate:

(1) The Probability $P(z_\ell|e_k)$ of obtaining a result z_ℓ from a given investigation e_k.

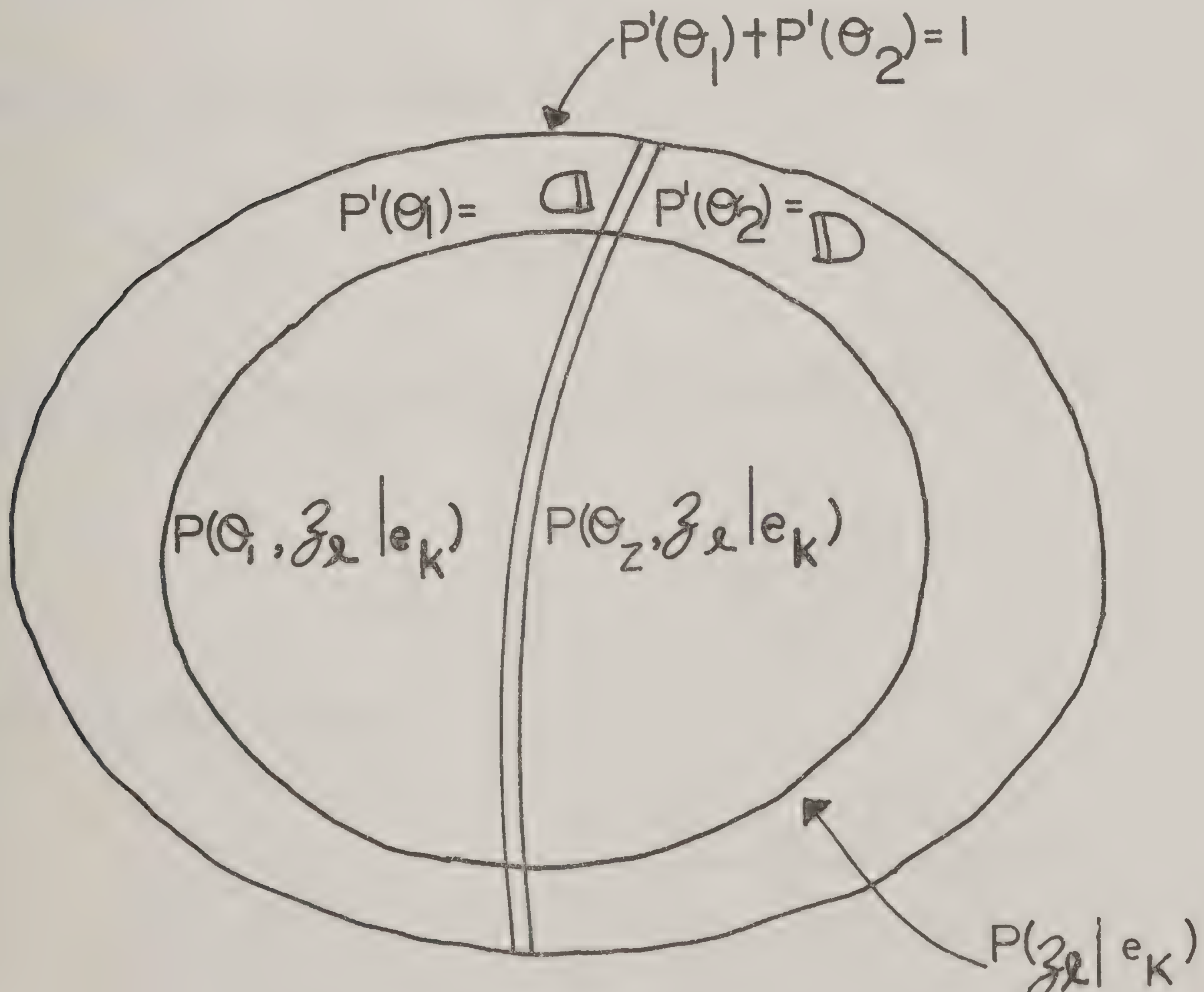
(2) The Posterior Probability $P''(S_i|z_\ell, e_k)$ of the existence of the true condition S_i when a given investigation e_k provides a given result z_ℓ.

Both of the above probabilities can be calculated from the available "judgment assigned" probability values.

$P'(S_i)$ = Prior probability of a true condition S_i.

$P(z_\ell|S_i, e_k)$ = Probability of obtaining a result z_ℓ from a given investigation e_k when a given true condition S_i exists.

The required probability calculation procedures are best presented by a Venn diagram. (See Reference (1), page 14)



The joint probability $P(S_i, z_\ell | e_k)$ of the occurrence of both the true condition S_i and the result z_ℓ of a given investigation e_k may be found by use of the basic definition of conditional probability, (See Reference (1), page 62),

$$P(z_\ell | S_i, e_k) = \frac{P(S_i, z_\ell | e_k)}{P'(S_i)}$$

which gives

$$P(S_i, z_\ell | e_k) = P(z_\ell | S_i, e_k) \cdot P'(S_i)$$

in terms of the available probability values.

From the Venn diagram the inner area which represents $P(z_\ell | e_k)$ is seen to be equal to the sum of all the joint probabilities,

$$P(S_1, z_\ell | e_k) + P(S_2, z_\ell | e_k) = \sum_{i=1}^{2} P(S_i, z_\ell | e_k)$$

thus, using the above equation for $P(S_i, z_\ell | e_k)$,

$$P(z_\ell | e_k) = \sum_{i=1}^{2} P(z_\ell | S_i, e_k) \cdot P'(S_i).$$

Now if we again employ the definition of conditional probability, the Posterior probability is given by

$$P''(S_i | z_\ell, e_k) = \frac{P(S_i, z_\ell | e_k)}{P(z_\ell | e_k)}$$

using the above results,

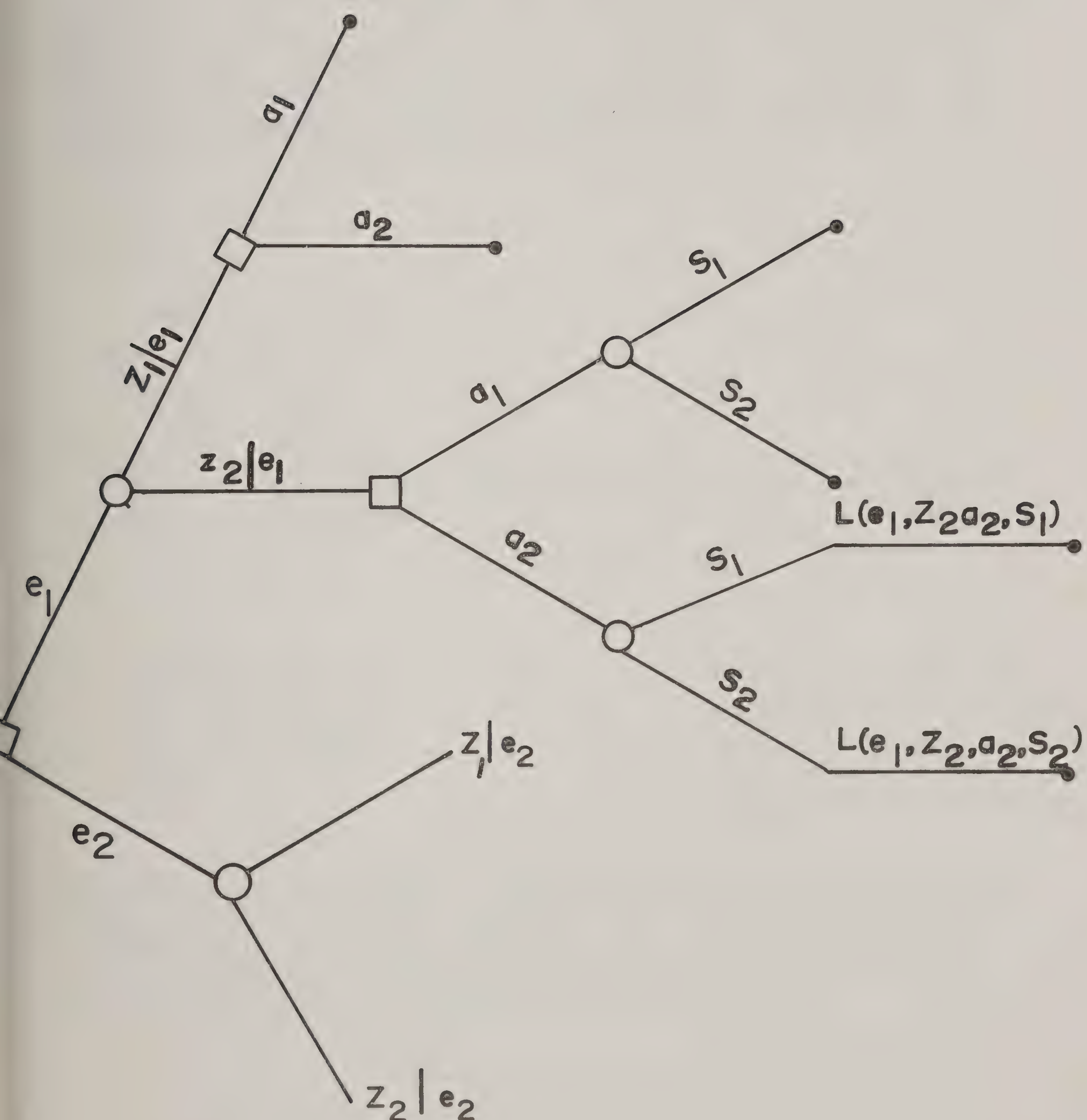
$$P''(S_i | z_\ell, e_k) = \frac{P(z_\ell | S_i, e_k) \cdot P'(S_i)}{\sum_{i=1}^{2} P(z_\ell | S_i, e_k) \cdot P'(S_i)}$$

which is in the form of Bayes' Theorem. (See Reference (1), page 119). For our example, above probability computations are organized as follows:

$P(z_\ell \mid S_i, e_k) \cdot P'(S_i)$	e_1		e_2	
	z_1	z_2	z_1	z_2
$P(z_\ell \mid S_1, e_k) \cdot P'(S_1)$	.28	.12	.36	.04
$P(z_\ell \mid S_2, e_k) \cdot P'(S_2)$	.18	.42	.06	.54
$P(z_\ell \mid e_k) = \sum_{\text{all } S_i}$	.46	.54	.42	.58
$P''(S_1 \mid e_k, z_\ell)$	.61	.22	.86	.07
$P''(S_2 \mid e_k, z_\ell)$	.39	.78	.14	.93

Selection of the Investigation with Minimum Average Loss by Working Backwards on the Decision Tree

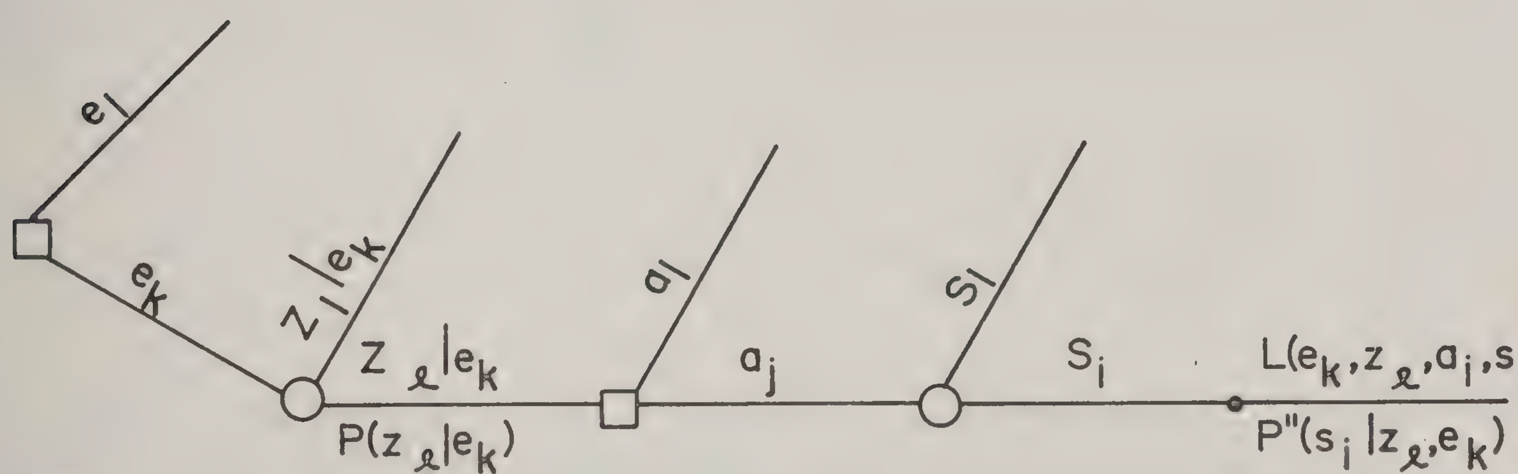
The decision tree display of the investigation selection problem is:



Each investigation e_k has its own respective set of result probabilities $P(z_\ell | e_k)$.

The result $z_\ell | e_k$ improves the knowledge concerning the true but unknown condition S_i, this improvement in knowledge being from the prior probability $P'(s_i)$ to the posterior probability $P''(S_i | z_\ell, e_k)$, which is used to select the decisions a_j.

Now since, before the selection of any investigation e_k, all information is known only on a probabilistic basis such as $P(z_\ell | e_k)$ and $P''(S_i | z_1, e_k)$, the value of an investigation e_k must be computed on an expected or average cost basis. Decisions concerning the selection of actions and investigations will be based on the minimum average loss or cost basis. This computation proceeds as follows, by working backwards from the action selection branch of the decision tree:



Position of Engineer when computing $\overline{L}(e_k)$ and finding minimum $\overline{L}(e^*)$

Position of Engineer when computing $\overline{L}$ (e_k, z_ℓ, a_j) and finding the a* which provides the minimum value $\overline{L}(e_k, z_\ell, a^*)$. Here, he imagines that $z_\ell | e_k$ is known.

For each of the combinations of investigations and corresponding results $z_\ell | e_k$, the engineer imagines that he has this information, and with the resulting posterior probability $P''(S_i | z_\ell, e_k)$, he positions himself at the action selection branch point of the decision tree. At this point, he is able to compute the average loss of each action a_j.

For each of the combinations of investigations and corresponding results $z_\ell|e_k$, the engineer imagines that he has this information, and with the resulting posterior probability $P''(S_i|z_\ell, e_k)$, he positions himself at the action selection branch point of the decision tree. At this point, he is able to compute the average loss of each action a_j,

$$\overline{L}(e_k, z_\ell, a_j) = \sum_{i=1}^{2} L(e_k, z_\ell, a_j, S_i) \cdot P''(S_i|z_\ell, e_k)$$

and select the action a* with the minimum average loss $\overline{L}(e_k, z_\ell, a^*)$. Thus for each of the possible investigation results $z_\ell|e_k$, there is a corresponding minimum average loss value $\overline{L}(e_k, z_\ell, a^*)$.

The engineer now moves to the investigation selection branch point of the decision tree and imagines that he has selected a given investigation e_k but has not yet obtained the results $z_\ell|e_k$. He does, however, know their probabilities $P(z_\ell|e_k)$, and the average loss or cost of e_k can be computed by

$$\overline{L}(e_k) = \sum_{\ell=1}^{2} \overline{L}(e_k, z_\ell, a^*) \cdot P(z_\ell|e_k)$$

This may be interpreted as being the average cost of all the optimum actions a* having minimum average loss, when investigation e_k is used. The engineer should select the investigation e* which has the minimum cost value $\overline{L}(e^*)$. For our example, the above average loss computations are organized as follows:

A. Compute $\overline{L}(e_k, z_\ell, a_j)$ and find minimum = $\overline{L}(e_k, z_\ell, a^*)$

Action a_1

$P''(S_i, e_k, z_\ell)$	e_1		e_2	
$L(e_k, z_\ell, a_1, S_i)$	z_1	z_2	z_1	z_2
$P''(S_1)$	.61	.22	.86	.07
$L(S_1)$	100	100	200	200
$L(S_1) \cdot P''(S_1)$	61	22	172	14
$P''(S_2)$	.39	.78	.14	.93
$L(S_2)$	500	500	600	600
$L(S_2) \cdot P''(S_2)$	195	390	84	558
$\sum_{\text{all } S_i} =$				
$\overline{L}(e_k, z_\ell, a_1)$	256	412	256	572
$\overline{L}(e_k, z_\ell, a^*) = \min \overline{L}(e_k, z_\ell, a_j)$	256	---	256	---

Action a_2

$P''(S_i \mid e_k, z_\ell)$	e_1		e_2	
$L(e_k, z_\ell, a_2, S_i)$	z_1	z_2	z_1	z_2
$P''(S_1)$	.61	.22	.86	.07
$L(S_1)$	300	300	400	400
$L(S_1) \cdot P''(S_1)$	183	66	344	28
$P''(S_2)$	.39	.78	.14	.93
$L(S_2)$	300	300	400	400
$L(S_2) \cdot P''(S_2)$	177	234	56	372
$\sum$ over all S_i =				
$\bar{L}(e_k, z_\ell, a_2)$	300*	300	400	400
$\bar{L}(e_k, z_\ell, a^*)$ =				
$\min \bar{L}(e_k, z_\ell, a_j)$	---	300	---	400

*These values are uniform because action a_2 provides adequate slide control for both conditions S_1 and S_2.

Summary for all Actions a_1, a_2

	e_1		e_2	
	z_1	z_2	z_1	z_2
$\bar{L}(e_k, z_\ell, a^*)$	256	300	256	400

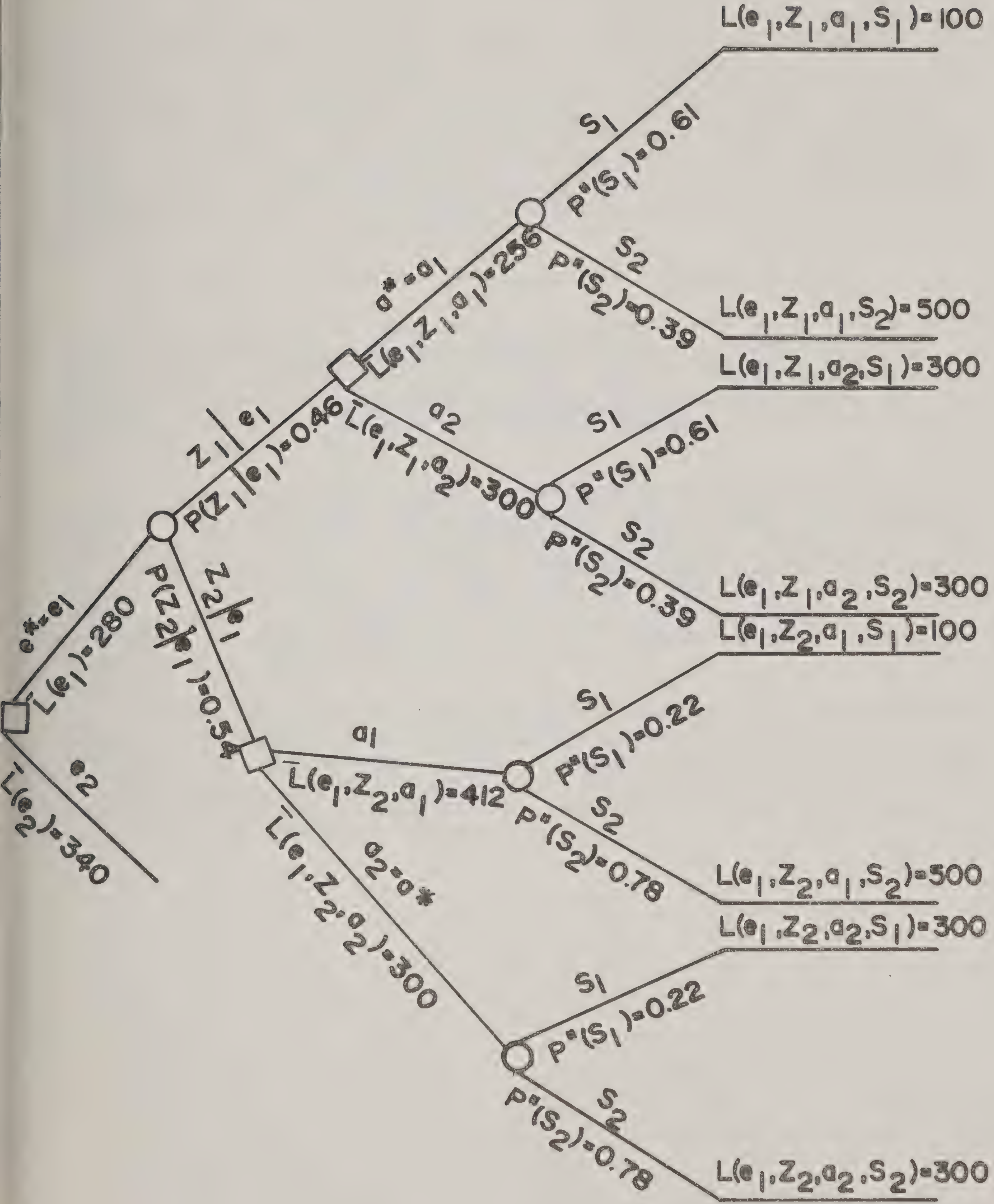
B. Compute $\bar{L}(e_k)$ and find $\bar{L}(e^*) = \text{min. of } \bar{L}(e_k)$,

where $\bar{L}(e_k) = \sum_{\ell=1}^{2} \bar{L}(z_\ell, e_k, a^*) \cdot P(z_\ell | e_k) = \sum_{\ell=1}^{2} \bar{L}(a^*) \cdot P(z_\ell)$

$P(z_\ell \mid e_k)$ $\bar{L}(e_k, z_\ell, a^*)$	e_1	e_2
$P(z_1 \mid e_k)$	.46	.42
$\bar{L}(e_k, z_1, a^*)$	256	256
$\bar{L}(a^*) \cdot P(z_1)$	118	108
$P(z_2 \mid e_k)$	.54	.58
$\bar{L}(e_k, z_2, a^*)$	300	400
$\bar{L}(a^*) \cdot P(z_2)$	162	232
$\sum_{\ell=1}^{2} \bar{L}(a^*) \cdot P(z_\ell)$	280	340

$\bar{L}(e^*) = \min \bar{L}(e_k) = 280 = \bar{L}(e_1)$

Therefore select e_1, the site investigation by the manual ditching method.



L(e1,Z1,a1,S1)=100
S1
P"(S1)=0.61
S2
P"(S2)=0.39
L(e1,Z1,a1,S2)=500
a*=a1
L̄(e1,Z1,a1)=256
L(e1,Z1,a2,S1)=300
a2
L̄(e1,Z1,a2)=300
S1
P"(S1)=0.61
S2
P"(S2)=0.39
L(e1,Z1,a2,S2)=300
Z1|e1
P(Z1|e1)=0.46
Z2|e1
P(Z2|e1)=0.54
e*=e1
L̄(e1)=280
e2
L̄(e2)=340
L(e1,Z2,a1,S1)=100
a1
L̄(e1,Z2,a1)=412
S1
P"(S1)=0.22
S2
P"(S2)=0.78
L(e1,Z2,a1,S2)=500
a2=a*
L̄(e1,Z2,a2)=300
L(e1,Z2,a2,S1)=300
S1
P"(S1)=0.22
S2
P"(S2)=0.78
L(e1,Z2,a2,S2)=300

C. POSTERIOR OR TERMINAL ANALYSIS, AFTER THE INVESTIGATION

Once having selected the optimum investigation e^*, the engineer orders its performance, and obtains the result z_ℓ. With this new information, he revises his prior probabilities $P'(S_i)$ to the posterior values $P''(S_i|z_\ell,e^*)$, and proceeds to select the best action a_j based on minimum expected loss calculations which employ the new revised $P''(S_i|z_\ell,e^*)$ values. It is useful to note that all necessary calculations for this selection of the best action have already been performed during the pre-posterior analysis selection of e^*; the expected loss $\overline{L}(e_k, z_\ell, a_j)$ based on the corresponding revised $P''(S_i|z_\ell, e_k)$ values for all possible combinations of e_k, z_ℓ, a_j are all available in the calculation tables.

For example, in terms of the roadway slope selection problem:

The engineer, having selected e_1, now performs his investigation and obtains the outcome z_ℓ, here for example: $z_\ell = z_1$, an indication of the upward strata condition S_1. He will select the action with the minimum average loss $\overline{L}(e_1, z_1, a^*)$ which is found in the computation tables as Action a_1 with an average loss value of 256. Therefore select action a_1, the 1 : 1 slope.

The entire decision sequence, as illustrated in this example, of selecting the investigation e_1, then obtaining the result z_1, and finally selecting the action a_1, is the best strategy that the engineer can employ in the face of uncertain conditions. It should be realized that his selection of action a_1 has a chance of being wrong, since there still remains a probability $P''(S_2|e_1, z_1) = 0.39$ of having the horizontal strata condition S_2 in the selected 1 : 1 slope of action a_1, with the corresponding slide problems and high loss $L(e_1, z_1, a_1, S_2) = \$500$.

However, without knowing the true strata condition S_i, the engineer has minimized his expected loss when he chooses a_1, and if he employs this strategy of minimum expected loss in all of his many professional decisions, his <u>total actual loss will be smaller</u> than if he employed some other strategy or no decision making strategy at all. Further developments and examples may be found in references (4) and (5).

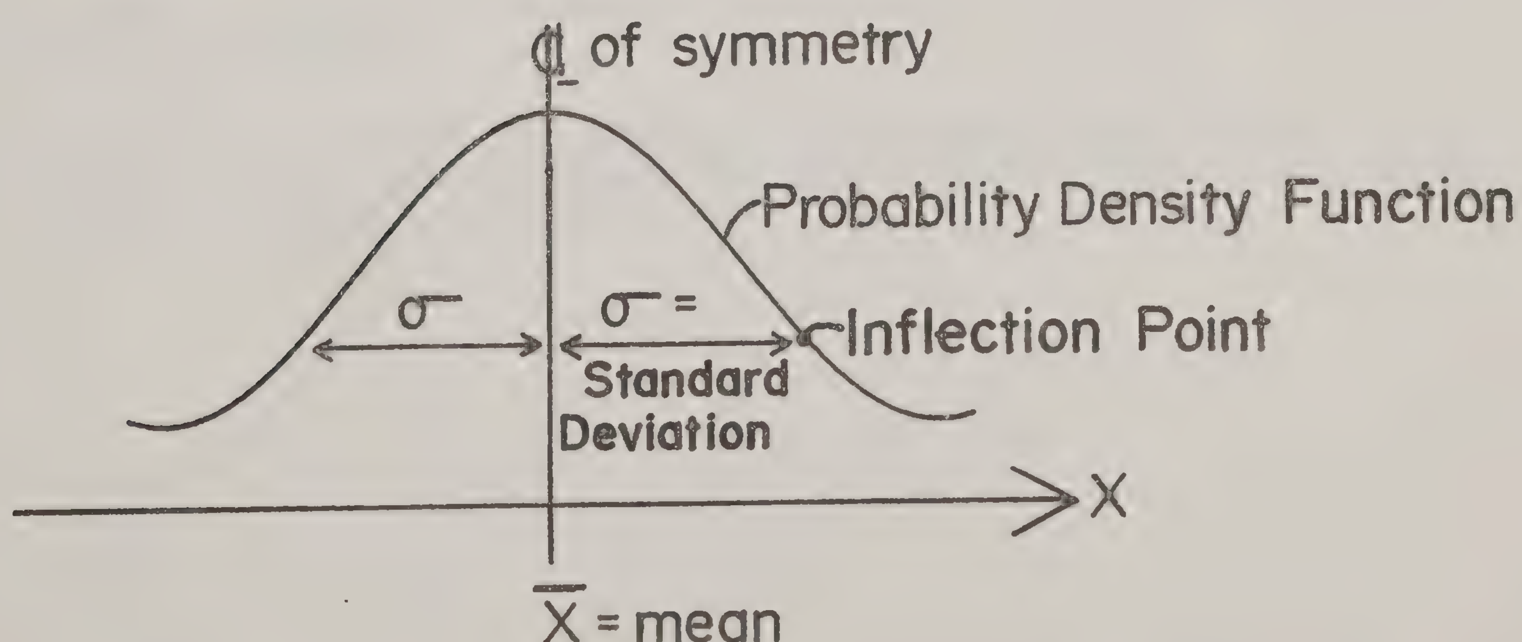
III. AN ADDITIONAL EXAMPLE

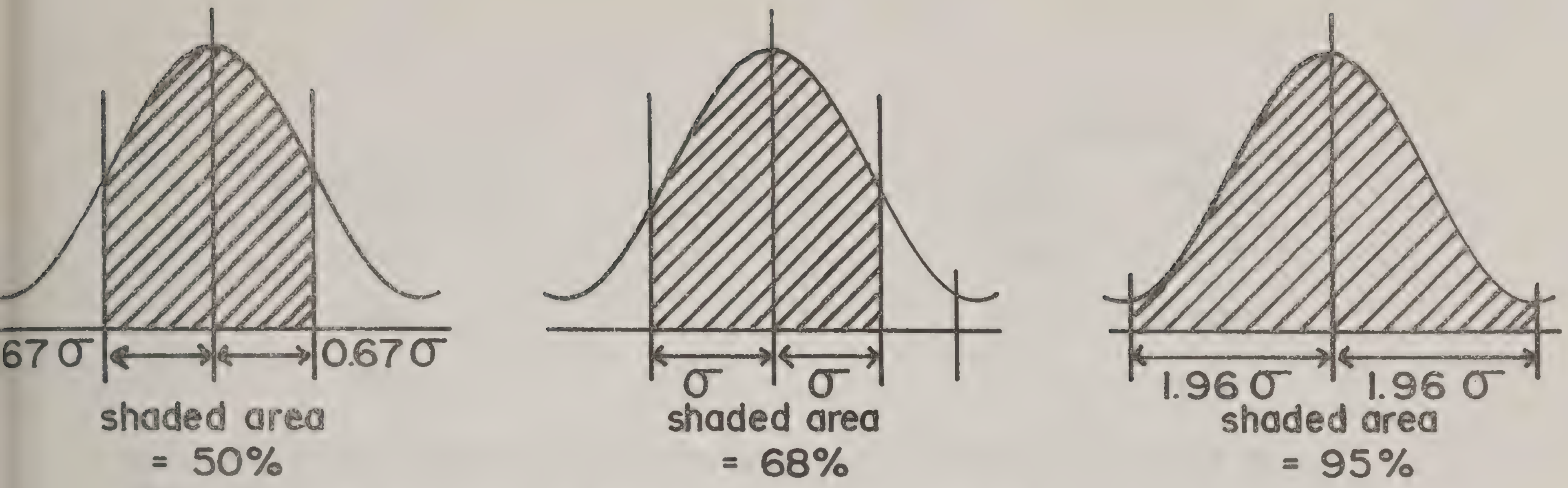
Terminal Analysis (Decision analysis after the sampling investigation)

Terminal analysis refers to the posterior analysis of the decision tree when new information about the conditions has been obtained by a sampling investigation. The optimum action is selected after the sampling investigation has been performed. In this type of analysis, the cost of sampling is considered to be an already incurred cost (sunk cost) and no longer enters into the problem.

The sampling investigation provides additional information concerning the unknown conditions and terminal analysis treats the problem of revising prior probabilities as a result of this sample information. The expected values of the various actions are computed using these revised probabilities (termed as posterior probabilities) to determine the best action.

The example in this section will deal with a normal probability distribution which is found to fit a number of common conditions in engineering decisions. Statement: The reader will perhaps not fully understand some of the terminology and procedures -- however, they are standard methods of engineering statistics. This example is given to show the powerful capabilities of the method to solve practical problems; lack of some procedural comprehension can easily be corrected by further study in standard texts. The normal distribution is fully tabulated in terms of the mean $\overline{X}$ (average value of the variable X) and a measure of dispersion or scatter of the individual values of the variable around the average value known as the standard deviation σ. The normal distribution has a probability density function which is a smooth, symmetric, continuous, bell-shaped curve as pictured in the figure below. The area under the curve over any interval on the horizontal axis represents the probability of the random variable, X, taking on a value in that interval. The area under the total curve is equal to unity. This curve reaches a maximum at the mean of the distribution. One half of the area lies on either side of the mean. The standard deviation σ controls the spread of the curve; for example, a large σ means a very wide curve. With any normal distribution, 0.50 of the area lies within $\pm$ 0.67 standard deviations from the mean; 0.68 of the area lies within $\pm$ 1.0 standard deviations; and 0.95 of the area lies within $\pm$ 1.96 standard deviations.





Example: An engineer is faced with the problem of accepting (action a_1) or rejecting (action a_2) a particular batch of Class "A" mix concrete delivered to an important job. The desired or specified mean concrete strength is 4000 psi. Assume that extensive observations on strengths of concrete of this Class "A" type have shown that the "population" strength X is normally distributed with a standard deviation of σ_X = 900 psi. The "population" here refers to the whole class of concrete about which conclusions are to be drawn. The engineer, based on his experience in concrete mix design, assigns the following prior probabilities of obtaining significantly different actual mean values S_i when class "A" mix is specified.

Actual Mean, S_i psi	Prior Probability $P'(S_i)$
3500	0.1
4000 (The desired value)	0.6
4500	0.3

Later, data from a sampling investigation will be used to revise or sharpen these prior probabilities. The data will consist of 9 cylinder tests with an arithmetic average value of

$$\overline{X}_o = 3700 \text{ psi}$$

this observed average is an estimate of the true mean S_i.

First however, we shall perform a "prior" decision analysis with only the prior probabilities $P'(S_i)$.

The engineer has studied the economic value problem extensively and has arrived at the following dollar losses for each of the conditions represented by the various possible actual means:

Actual Mean S_i	Loss, \$ a_1 Accept	a_2 Reject
3500	\$1000	\$ 0
4000	\$ 0	\$ 500
4500	\$ 0	\$1000

Using the prior probabilities, the expected loss can be computed as shown below.

Actual Mean S_i	Loss Accept		Reject		$P'(S_i)$		Prior Expected Loss Accept	Reject
3500	\$1000	or	\$ 0	x	0.1	=	\$100	\$ 0
4000	\$ 0	or	\$ 500	x	0.6	=	\$ 0	\$300
4500	\$ 0	or	\$1000	x	0.3	=	\$ 0	\$300
					Total expected loss		\$100	\$600

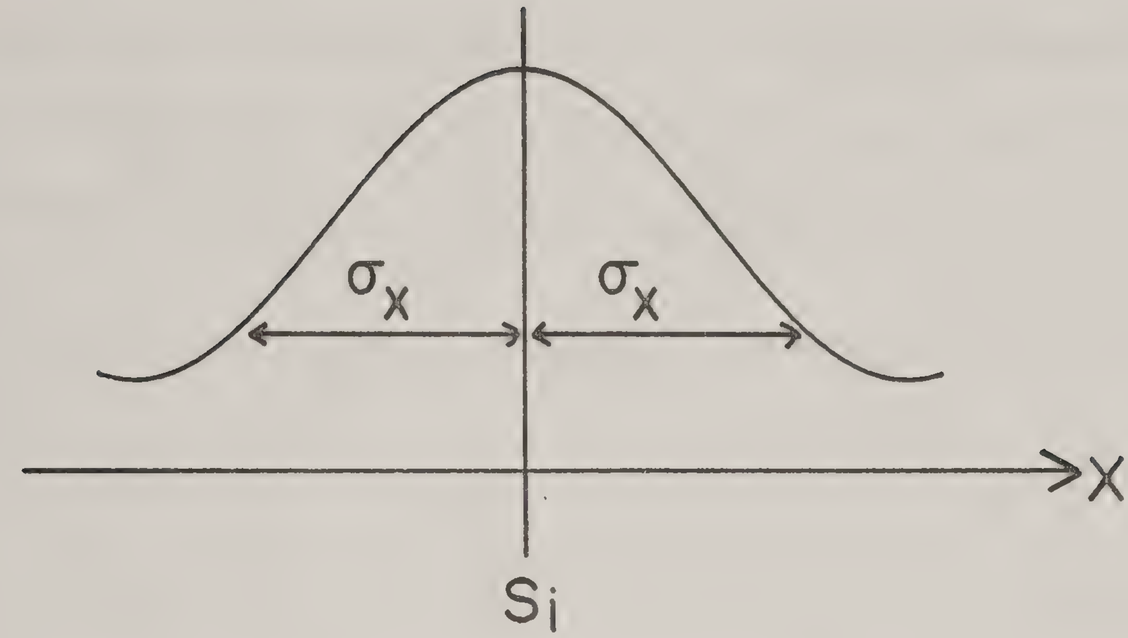
It can be seen that if the sample data is not used, the prior expected losses of accepting, or rejecting are \$100 and \$600, respectively. Thus, using prior probabilities only, the mix should be accepted. It is assumed that concrete cylinder strengths X follow a normal distribution; the parameters which "sufficiently" describe this process are the sample mean X and the number of samples n. For large samples the distribution of sample means $\overline{X}$ is also normal and the standard deviation of the sample means from the true mean is given by

$$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}}$$

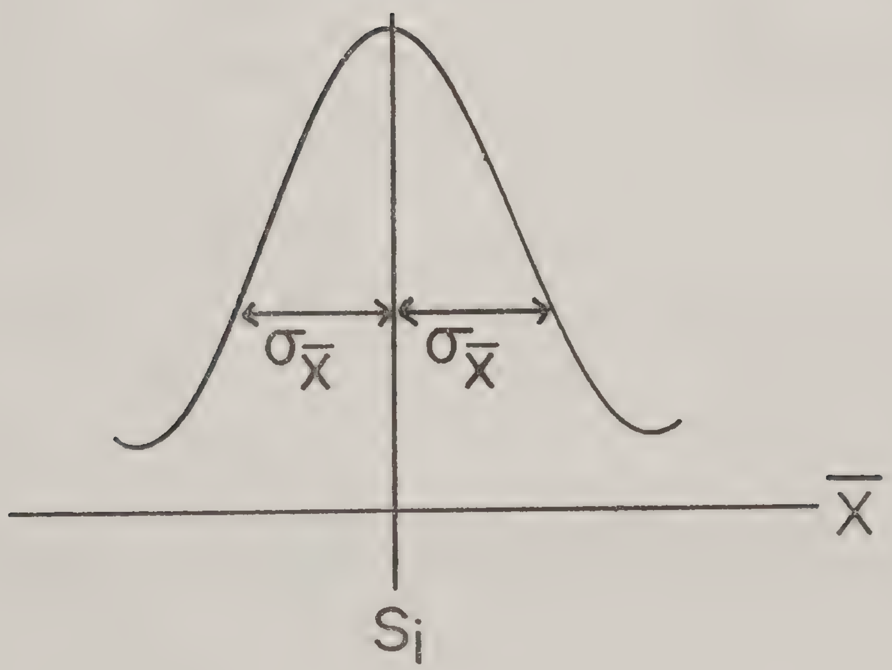
where

σ_X = standard deviation of the cylinder strength X

n = number of observations of X = size of sample



Normal Distribution of Cylinder Strength
about a possible actual mean value S_i

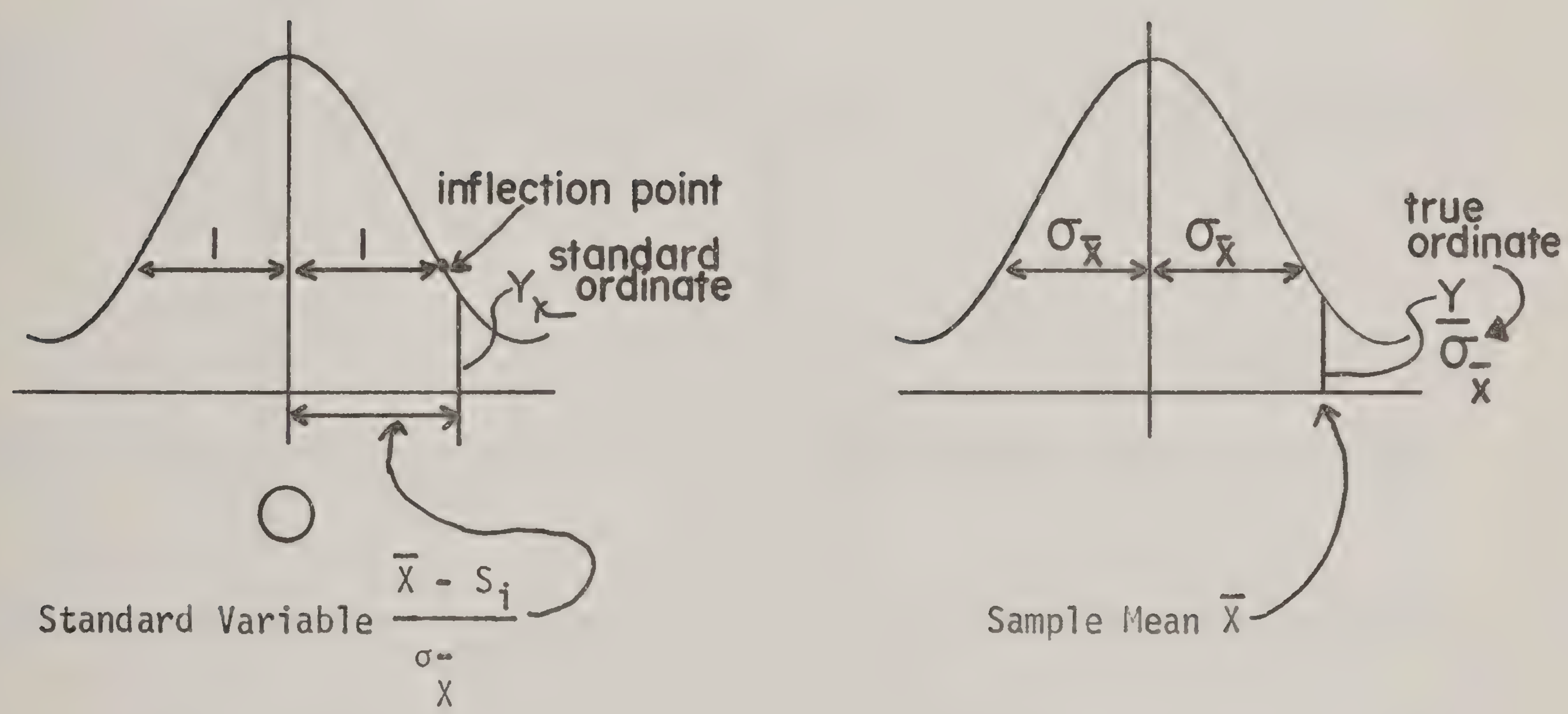


Normal Distribution of a Sample Mean $\overline{X}$
equal to the arithmetic average of n cylinder tests X

In this example:

$$\sigma_{\bar{X}} = \frac{900}{\sqrt{9}} = 300 \text{ psi}$$

The sample likelihood* or the probability of obtaining the observed sample mean $\bar{X}$ as a function of the true mean S_i is computed for each possible value of the true means S_i. First $\frac{\bar{X} - S_i}{\sigma_{\bar{X}}}$ is calculated, then using this value, the ordinate of the normal curve is read from the tables of the normal distribution. (a sample is attached).



* See Benjamin and Cornell, "Statistics and Probability for Civil Engineers," McGraw-Hill, for terminology and notation used in the following calculations.

Ordinates of the Normal Curve

$$y = \frac{1}{2\pi} e^{\frac{-z^2}{2}}$$

z	.00		.07	.09
.0				
.6			.3187	
1.0	.2420			
2.6			.0113	
3.9				

The above ordinates give the probability density for $z = (x-\mu)/\sigma$ deviations from the mean μ (that is, $z = 0$). To fit a normal frequency curve to observed data consisting of n observations, multiply the ordinate from the table for any value of z by n/σ. To fit a normal probability curve multiply the ordinate by $1/\sigma$.

Table A-3 "Basic Statistical Methods for Engineers and Scientists," Neville and Kennedy, International Textbook Company.

Condition S_i	standard variable $\frac{\overline{X} - S_i}{\sigma_{\overline{X}}}$	standard ordinate y (from Tables)	P.D.F. True ordinate $= \frac{y}{\sigma_{\overline{X}}}$
3500	$\frac{3700 - 3500}{300} = 0.67$	0.3187	$\frac{0.3187}{300}$
4000	$\frac{3700 - 4000}{300} = -1.00$	0.242	$\frac{0.242}{300}$
4500	$\frac{3700 - 4500}{300} = -2.67$	0.0113	$\frac{0.0113}{300}$

Condition S_i	Prior $P'(S_i)$		Sample likelihood $\ell(\overline{X},n\mid S_i)$		Joint Measure $\ell(\overline{X},n,S_i\mid e)$	Posterior Prob. $P''(S_i\mid\overline{X},n,e)$
3500	0.1	x	$\frac{0.3187}{300}$	=	$\frac{0.03187}{300}$	0.177
4000	0.6	x	$\frac{0.242}{300}$	=	$\frac{0.14520}{300}$	0.805
4500	0.3	x	$\frac{0.0113}{300}$	=	$\frac{0.00330}{300}$	0.018
				Σ =	$\frac{0.18043}{300}$	1.000

The sample likelihoods computed above are the relative probabilities of finding the actual sample mean given that the mean is one of three possible values 3500, 4000, and 4500 psi.

The posterior probabilities are calculated by the following expression:

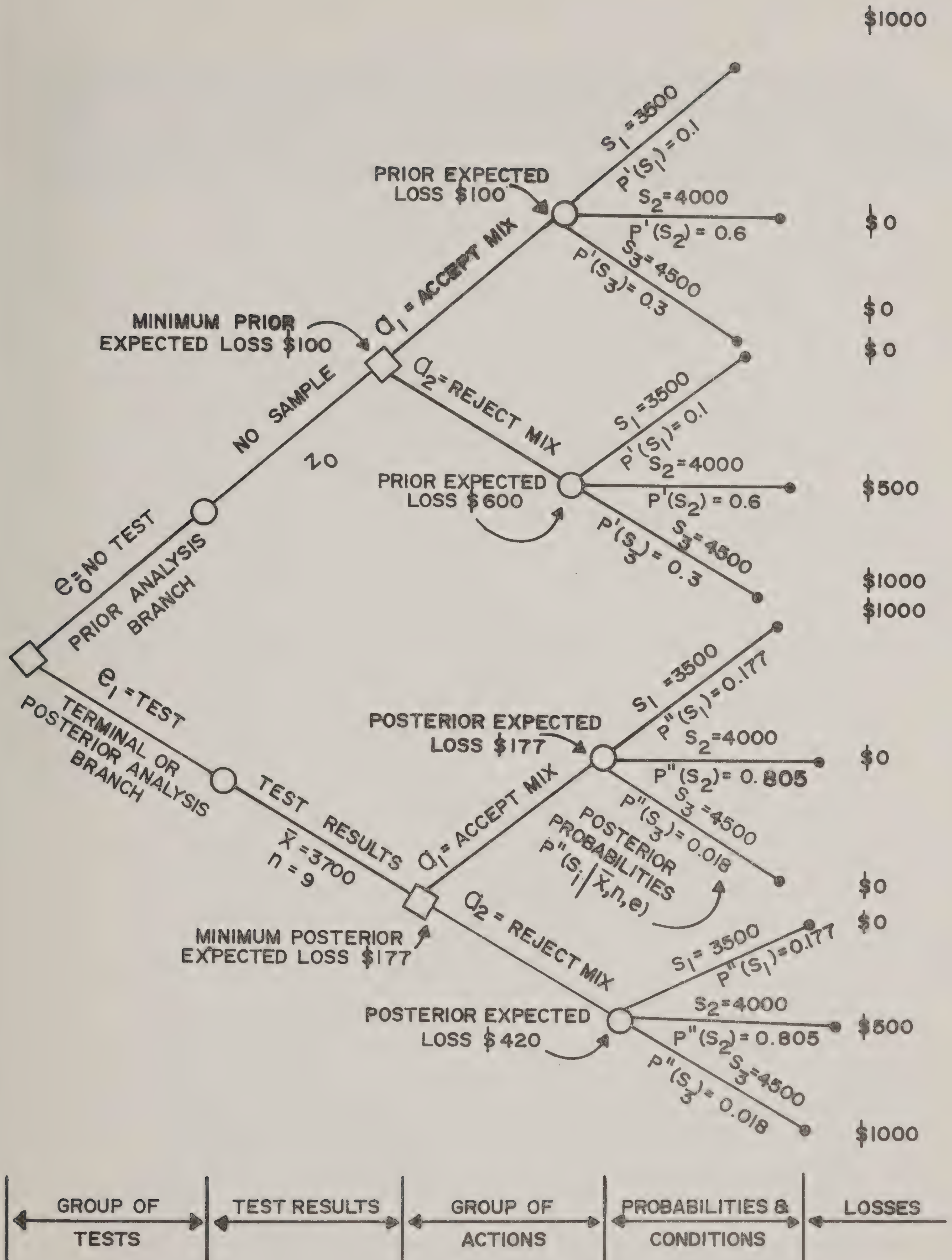
$$P''(S_i|z) = P'(S_i)\cdot\ell(z|S_i)\cdot n(z) \text{ , where } z = \text{sample information } (\overline{X},n)$$

or

Posterior Probability = Prior Prob. times sample likelihood times a normalizing constant

The joint measures are calculated by multiplying the prior probabilities of state by the sample likelihood given that particular state. The joint measures are then summed to give the reciprocal of the normalizing constant. The normalizing constant is introduced so that the sum of the posterior probabilities, for all of the conditions considered, will equal unity. The posterior probabilities then are calculated by taking the ratio of each of the joint measures to their sum. Terminal analysis concludes with the computation of the expected posterior losses as shown in the following table:

Condition	Loss, $		Posterior Probability	Posterior Expected Loss	
S_i	Accept	Reject	$P''(S_i\|\overline{X},n,e)$	Accept	Reject
3500	$1000	$0	0.177	$177	$0
4000	$0	$500	0.805	$0	$402
4500	$0	$1000	0.018	$0	$ 18
Total expected posterior loss			=	$177	$420



TERMINAL ANALYSIS DECISION TREE

CONCRETE MIX EXAMPLE

It should be noted that the posterior probabilities differ from the prior probabilities by virtue of the sample results. The sample mean strength was lower than the desired mean strength; this produces an increase in the expected cost of acceptance of the mix. It should also be noted that the limitation of S_i to three specific values restricts the posterior probabilities to those same values.

REFERENCES

(1) Parzen, E.; "Modern Probability Theory and Its Applications"; John Wiley and Sons, New York

(2) Lieberman, and Bowker; "Engineering Statistics"; Prentice-Hall, Inc., New Jersey

(3) Kissam, Phillip; "Surveying for Civil Engineers"; McGraw-Hill, New York

(4) Benjamin, and Cornell; "Statistics and Probability for Civil Engineers"; McGraw-Hill, New York

(5) Raiffa; "Decision Analysis"; Addison-Wesley, Reading, Massachusetts

(6) Neville and Kennedy; "Basic Statistical Methods for Engineers and Scientists"; International Textbook Co., Scranton, Pa.